UNIVERSITY PHYSICS SERIES.

A Group of Text Books for Intermediate and Advanced Courses.

LEIGH PAGE, PH.D.

CONSULTING EDITOR.

Introduction to Physical Optics. *By* John K. Robertson, F.R.S.C., *Professor of Physics, Queen's University.*

Principles of Electricity. *By* Leigh Page, Ph.D., *Professor of Mathematical Physics, Yale University;* and Norman Ilsley Adams, Jr., Ph.D., *Assistant Professor of Physics, Yale University.*

Physical Mechanics. *By* Robert B. Lindsay, Ph.D., *Associate Professor of Theoretical Physics, Brown University.*

PRINCIPLES OF
Electricity

An Intermediate Text in Electricity and Magnetism

BY

LEIGH PAGE, PH.D.
Professor of Mathematical Physics in Yale University

AND

NORMAN ILSLEY ADAMS, JR., PH.D.
Associate Professor of Physics in Yale University

SIXTH PRINTING

NEW YORK
D. VAN NOSTRAND COMPANY, INC.
250 FOURTH AVENUE

First Published, June 1931

*Reprinted, July 1932, August 1934, September 1936
August 1939, July 1940*

Printed in U. S. A.

PRESS OF
BRAUNWORTH & COMPANY, INC.
BUILDERS OF BOOKS
BRIDGEPORT • CONNECTICUT

PREFACE

The purpose of the authors in preparing this book is to provide an elementary text suitable both for undergraduates who have completed general courses in physics and in the calculus such as are usually given in Sophomore year, and for advanced students who wish a more comprehensive training in the important subject of electromagnetism. By utilizing the first twelve chapters with the possible exception of Chapter III, a course of the first type of the usual length may be arranged. In order to make it possible to take up magnetostatics before electrostatics, the more fundamental deductions of Chapters I and II have been repeated in Chapter IV.

Considerable space has been devoted to the theory of electrical measurements in the belief that a didactic course should be accompanied by practice in the laboratory. Nevertheless, the main emphasis has been placed, not on applications of the theory, but on instilling correct concepts in the mind of the reader. While the relativity principle — so intimately connected with electromagnetism — is not mentioned *per se*, the point of view of that theory is maintained throughout. Needless to say, the electron theory of matter is adopted from the beginning. More space than usual has been devoted to the phenomena of high frequency circuits and electromagnetic waves, matters of great interest to everyone in an age of electrical communication. Every effort has been made to develop the subject in a logical manner and not to slight important deductions in order to minimize analysis. Carefully selected problems have been appended to most of the articles.

L. P.
N. I. A., Jr.

YALE UNIVERSITY,
February 1931.

CONTENTS

CHAPTER I

Fundamental Laws of Electrostatics

CHAPTER II

Dielectrics and Conductors

CHAPTER III

Solution of Electrostatic Problems

CHAPTER XV

Coupled Circuits, Filters and Lines

CHAPTER XVI

Electromagnetic Waves

CHAPTER XVII

High Frequency Oscillations

CONVERSION TABLE OF ELECTRIC AND MAGNETIC QUANTITIES

A quantity as measured in electromagnetic units (e.m.u.) is indicated by a letter with the subscript m, in electrostatic units (e.s.u.) by the same letter with the subscript s, and in practical units by the subscript p. The unit, where it has received a name, is given in parentheses.

QUANTITY	E.M.U.	E.S.U.	PRACTICAL
Distance	l_m (cm)	$= l_s$ (cm)	$= (10)^9 l_p$
Time	t_m (sec)	$= t_s$ (sec)	$= t_p$ (sec)
Mass	m_m (gm)	$= m_s$ (gm)	$= (10)^{-11} m_p$
Force	F_m (dyne)	$= F_s$ (dyne)	$= (10)^{-2} F_p$
Work or Energy	U_m (erg)	$= U_s$ (erg)	$= (10)^7 U_p$ (joule)
Power	$\mathcal{P}_m$	$= \mathcal{P}_s$	$= (10)^7 \mathcal{P}_p$ (watt)
Charge	Q_m	$= \frac{1}{3}(10)^{-10} Q_s$	$= (10)^{-1} Q_p$ (coulomb)
Electric Intensity	E_m	$= 3(10)^{10} E_s$	
Polarization	P_m	$= \frac{1}{3}(10)^{-10} P_s$	
Displacement	D_m	$= \frac{1}{3}(10)^{-10} D_s$	
Magnetic Intensity	H_m (gauss)	$= \frac{1}{3}(10)^{-10} H_s$	
Intensity of Magnetization	I_m	$= 3(10)^{10} I_s$	
Magnetic Induction	B_m (gauss)	$= 3(10)^{10} B_s$	
Potential or e.m.f.	V_m	$= 3(10)^{10} V_s$	$= (10)^8 V_p$ (volt)
Current	i_m	$= \frac{1}{3}(10)^{-10} i_s$	$= (10)^{-1} i_p$ (ampere)
Resistance	R_m	$= 9(10)^{20} R_s$	$= (10)^9 R_p$ (ohm)
Capacity	C_m	$= \frac{1}{9}(10)^{-20} C_s$	$= (10)^{-9} C_p$ (farad)
Inductance	L_m	$= 9(10)^{20} L_s$	$= (10)^9 L_p$ (henry)

Example: Find the magnitude of a current of 50 amp in e.m.u. and e.s.u.

$$i_m = (10)^{-1} i_p = 5 \text{ e.m.u.},$$
$$i_s = 3(10)^9 i_p = 1.5(10)^{11} \text{ e.s.u.}$$

All equations in electromagnetic units hold also for practical units, but the form of many equations differs when we pass from electromagnetic units to electrostatic units. See article 108. As κ and μ are dimensionless ratios they are the same in all units.

Electric intensity is often expressed in the hybrid unit volt/cm, this being the field present when the potential drops one volt in a distance of one centimeter. Denoting the electric intensity in terms of this unit by E_p,

Electric Intensity........................ $E_m = 3(10)^{10} E_s = (10)^8 E_p$ (volt/cm)

The following prefixes have the indicated significance when attached to a unit:

mega = 1 million $[(10)^6]$,
kilo = 1 thousand $[(10)^3]$,
milli = 1 thousandth $[(10)^{-3}]$,
micro = 1 millionth $[(10)^{-6}]$.

Example: 1 megohm = $(10)^6$ ohm, 1 micromicrofarad = $(10)^{-12}$ farad.

CHAPTER I

FUNDAMENTAL LAWS OF ELECTROSTATICS

1. Attraction and Repulsion. — The discovery of electrification, which appears to have been made by the Greeks about 600 B.C., consisted in the observation that a piece of amber which has been rubbed acquires the property of attracting light bodies to itself. It was not until 1600 that it was noticed that electrified bodies may repel as well as attract, and du Fay, about 1735, was the first to appreciate the fact that there are two distinct kinds of electricity.

These two kinds of electricity are called *positive* and *negative*. When a glass rod is rubbed with silk the glass acquires a positive (*vitreous*) charge of electricity and the silk a negative charge. Conversely, when an ebonite rod is rubbed with fur, the ebonite obtains a negative (*resinous*) charge and the fur a positive charge. The recognition of two kinds of electricity rests upon the observation that two electrified glass rods, or two electrified ebonite rods, repel each other, whereas an electrified glass rod attracts an electrified ebonite rod. The results of early experiments may be embodied in the following qualitative laws:

(I) *Like charges repel, unlike charges attract.*

(II) *The force between two charges decreases as the distance between them is increased.*

We have noted that when a glass rod is rubbed with silk, the silk is electrified as well as the glass, and with a charge of opposite sign to that on the glass. This fact suggests that electrification consists in the *separation* of equal charges of opposite sign rather than in the creation of charges. In more detail it indicates that the glass and the silk in their original unelectrified condition each contains equal amounts of positive and negative electricity, the attraction of the one just sufficing to neutralize the repulsion of the other on any third charge, and that when the two are rubbed

together negative electricity is transferred from the glass to the silk, or positive electricity from the silk to the glass, or both. If, now, we keep the silk tightly wound around the glass rod after the two have been rubbed together we find that the combination exhibits no evidence of electrification, showing that the negative charge on the former neutralizes the positive charge on the latter and therefore is equal to it in magnitude. Thus we confirm our surmise that electrification consists in a separation and not in a creation of charges.

While ideally we might suppose that continued rubbing of glass with silk would result in a transfer of all the positive electricity in the combination to the glass and of all the negative electricity to the silk, it is found that this limiting state is far from being reached by any methods available in the laboratory. In fact, our present knowledge of the constitution of matter indicates that the most intense electrification obtainable in gross matter represents the transfer of a very minute portion of the electricity originally contained in the bodies concerned.

Furthermore, as will appear from the results of experiments to be described later, electricity is not infinitely divisible, but consists of small discrete entities. The elementary negatively charged particle, known as the *electron*, has a charge of $-4.77(10)^{-10}$ electrostatic unit and a mass of $9.0(10)^{-28}$ gram. All negative charges consist of an integral number of electrons. The charge of the electron is so small, however, that ordinary electrical experiments fail to reveal the fact that the least amount by which two negative charges may differ is the charge of a single electron.

The elementary positively charged particle is known as the *proton*. It has a charge of $4.77(10)^{-10}$ electrostatic unit and a mass of $1.66(10)^{-24}$ gram. The charge of the proton is equal and opposite in sign to that of the electron, but its mass, which is substantially the same as that of the hydrogen atom, is 1850 times greater. Therefore the proton is much less mobile than the electron. We may consider that atoms are built up of protons and electrons and nothing else. The hydrogen atom consists of a

single proton and a single electron, whereas the heavier atoms contain a minute nucleus built up of four or more protons together with a smaller number of electrons around which are distributed enough additional electrons so that in its normal state the total negative charge of the atom is equal to its total positive charge.

The fact that the electron is much lighter and therefore much more mobile than the proton indicates that electrification consists solely in the transfer of negative electricity. So great is the evidence supporting this view that we may conceive of the positive atomic nuclei as fixed in the bodies of which they are constituents, and of electrification as due to an excess or defect of electrons. Thus when glass is rubbed with silk the contact of the two dissimilar substances results in the detachment of electrons from surface atoms in the glass and their attachment to surface atoms in the silk. The glass becomes positively charged as a result of the subtraction of electrons, and the silk negatively charged in consequence of the addition of electrons.

2. Insulators and Conductors. — With respect to their electrical properties solids may be classified as *insulators* or *conductors*. The electrons in the former are so tightly attached to the atoms that only on the surface can they be torn loose and here only by relatively large forces. Insulators, or *dielectrics* as we shall more frequently call them, are characterized by the fact that thin sheets are transparent or at least translucent to light. If we electrify an insulator, such as a glass rod, by tearing off electrons from a few spots on the surface, the remaining electrons in the rod are unable to redistribute themselves under the action of the electrical forces to which the electrification gives rise. Therefore the charges remain at the spots on the surface at which they have been produced. Nevertheless a dielectric does not remain altogether uninfluenced by electrical forces. While the electrons in each atom cannot easily be completely detached from the atom, they suffer a small displacement from their equilibrium positions under the action of electrical forces. Although the displacement of the electrons in a single atom is too small to produce noticeable

effects, a displacement of the electrons in all the atoms of an insulator in the same direction gives rise to pronounced effects which are of great importance in many electrostatic phenomena.

In conductors, on the other hand, the outer electrons normally present in each atom become loosened from the atom and are free to wander through the solid under the influence of electrical forces. Therefore, if a charge is produced on the surface of a conductor by rubbing, these *free electrons* redistribute themselves under the forces of electrical attraction and repulsion with the result that the charge does not necessarily remain localized in its original situation. As the human body is a fair conductor, charges produced on a conducting rod held in the hand do not remain on the rod, but run over to the hand and through the experimenter's body to earth. If it is desired to retain the charge on the conducting rod it must be supported by an insulating handle. As a class conductors are opaque to light except in the very thinnest sheets. Metals are the best conductors of electricity, particularly such metals as copper and silver. As will be shown later, good conductors of electricity are also good conductors of heat, and *vice versa*.

If two conductors are placed in contact the free electrons in the one have no great difficulty in passing across the boundary surface into the other. On the other hand electrons do not easily escape through the surface of a conductor into an adjacent gas or vacuum. Only under especial conditions, such as when enormous electric fields are applied or high temperatures used, or when ultra-violet light or X-rays are allowed to impinge on a metal, are electrons emitted from its surface into the surrounding space.

Next we must mention an important class of liquid conductors known as *electrolytes*. In these the charged particles which are free to move under the influence of the electrical forces are not electrons, but charged atoms or groups of atoms, known as *ions*. The carriers of electricity are of both signs, the positive ions moving in one direction under the action of electrical forces and the negative ions in the opposite. The most important electrolytes are solutions of salts or acids in water.

Finally, a gas may be made conducting by the passage through it of X-rays or rays from radioactive substances, or by high temperature such as exists in a flame. In this case electrons are torn loose from normally uncharged atoms or molecules and either move freely or attach themselves to other molecules. Therefore we have present in the gas ions of both signs which move in opposite directions under the influence of electrical forces until they become neutralized by recombination or reach a metal electrode.

3. Electrification by Induction. — Once a positive charge has been produced on the end of a glass rod by rubbing it with silk any number of additional charges can be produced on nearby conductors by the process of *induction.* Consider the uncharged metal rod *AB* (Fig. 1) with rounded ends supported by the insulating stand *C.* If the charged end *G* of the glass rod is brought near to the end *A* of the metal, the electrons in the conductor are attracted to *A*, giving rise to a negative charge on this end of the metal rod and leaving the end *B* positively charged. The charges so produced are said to be *induced.* On removal of the glass rod the metal resumes its original unelectrified condition. If, however, the experimenter touches the metal rod at *B* while *G* is held in the position indicated in the figure, additional electrons pass up from the ground through the observer's body to *AB* under the attraction of the positive charge on the glass rod. If, now, the experimenter removes his finger from contact with *B* while still keeping *G* close to *A*, the metal rod retains the excess electrons, thus acquiring a net negative charge. On final removal of the glass rod this negative charge spreads over the whole of the surface of *AB.*

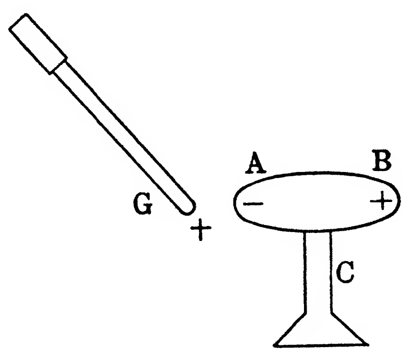


Fig. 1

By a slight modification of this procedure charges of both signs can be induced. All that is necessary is to replace the metal rod by two uncharged metal spheres *A* and *B* (Fig. 2) suspended by insulating silk threads. The spheres are held in contact with each other and the positively electrified glass rod *G* is brought near to *A*. Under the attraction of the charge on *G* electrons pass from *B* to *A*. If, now, *A* and *B* are separated before *G* is removed, the first is left with a negative charge and the second with an *equal* positive charge. Note that in the former experiment the conductor consisting of the observer's body and the earth took the place of the sphere *B* and was left with a positive charge at the end of the process.

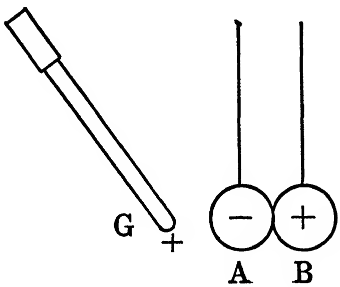


FIG. 2

Electrostatic induction is responsible for the attraction exerted by a charge on uncharged bodies. On account of its closer proximity to *G* the negative charge on the end *A* of the metal rod in Fig. 1 is attracted by the positive charge on the glass rod more than the equal positive charge at *B* is repelled. Consequently the rod *AB*, if free to move, would be drawn toward *G*. If the rod *AB* were a dielectric instead of a conductor, the electrons bound to the atoms of which it is composed would be slightly displaced to the left by the attraction of the charge at *G*, giving rise to a negative charge on the end *A* and a positive charge on the end *B*. As the electrons are tightly fastened to the atoms in this case the charge cannot be altered to any considerable extent by touching the rod with the finger, but the rod is attracted by the charge at *G* for precisely the same reason as in the case of a conductor.

Problem 3a. Does it matter if the experimenter touches *A* rather than *B* in inducing a charge on the metal rod of Fig. 1? Give reasons.

4. The Electroscope. — This simple instrument is very useful in investigating electrostatic phenomena in a qualitative manner, and when calibrated it can be used for quantitative measurements. Many forms of electroscope have been designed, but we shall describe only the simplest type consisting of a pair of gold leaves *G*, *G* (Fig. 3) suspended from a metal rod *R* which supports a conducting disk *D* on its upper end. The metal parts are insulated from the ground by the glass bottle *B* through the neck of which the rod *R* projects. On the disk *D* may be placed a metal pail *P*.

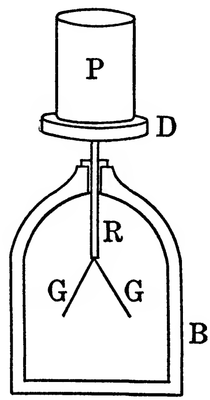


FIG. 3

With this instrument Faraday performed the famous "ice-pail" experiment, so named because he used a common metal ice-pail for *P*. Lowering the positively charged sphere *B* of Fig. 2 into *P* without touching it to the sides or bottom of the pail, electrons were attracted to the inner surface of *P* by the positive charge on the sphere, leaving the gold leaves positively electrified. Having like charges the gold leaves repelled each other and therefore diverged. The following phenomena were noted:

(*a*) Provided the sphere is well below the mouth of the pail the degree of divergence of the gold leaves is independent of its position.

(*b*) If the charged sphere is brought into contact with the bottom of the pail the divergence of the gold leaves remains unchanged, and does not decrease when the sphere is completely removed. The sphere, however, is found to have lost its entire charge.

This experiment shows that the charge induced on the inner walls of the pail is equal in magnitude though opposite in sign to the charge on the sphere. When the sphere is touched to the pail its entire charge escapes to the pail, neutralizing the negative charge on the inner walls of the latter. Since no charge can be acquired by again bringing the discharged sphere into contact

with the inside of the pail, there can be no charge left on its inner walls. The positive charge originally on the sphere now resides entirely on the gold leaves, rod and outer walls of the pail.

If the electroscope is charged positively, the introduction of a positive charge into the pail P causes the divergence of the gold leaves to increase, for electrons are attracted to the inner surface of the pail, leaving the leaves with an increased positive charge. On the other hand, the introduction of a negative charge into the pail causes electrons to be repelled from the pail to the gold leaves, diminishing the positive charge on the latter and decreasing the divergence. If the negative charge introduced into the pail is large enough, sufficient electrons may be repelled to the leaves to cause them to collapse as the charge approaches the pail and then to diverge again with a negative charge. In this way the electroscope may be used to determine the sign of the charge on an electrified body, or the sign of the charge on the gold leaves may be ascertained if that of the charge on the body introduced into the pail is known.

If two uncharged rods, one of glass and the other of ebonite, are introduced into the pail of an uncharged electroscope and rubbed together inside the pail, no divergence of the gold leaves is noted. If either rod is removed, however, the gold leaves diverge with a charge of the same sign as that of the rod remaining in the pail. This experiment constitutes a more precise method than that previously described of showing that the negative charge acquired by the ebonite rod is exactly equal to the positive charge obtained by the glass and therefore that electrification consists solely in the separation of charges already in existence.

The experiments described above enable us to draw the following conclusions regarding the relative magnitudes of charges:

(*a*) Two charges of opposite sign are of the same magnitude if they produce no effect on the gold leaves when they are placed together inside the pail of the electroscope.

(*b*) Two charges, A and B, of like sign are of equal magnitude if the one produces the same effect on the gold leaves as the other when introduced into the pail.

(*c*) A charge *C* has twice the magnitude of *A* if it produces the same effect on the gold leaves as *A* and *B* together.

5. Induction Machines. — The simplest type of induction machine is the *electrophorus*. It consists of a disk of sealing wax *AB* (Fig. 4) the upper surface of which has been charged negatively by being rubbed with fur. Above the sealing wax base is a metal disk *MN* of the same area supported by an insulating handle *H*. If the experimenter's finger is touched for a moment to *MN* while in the position indicated in the figure, electrons are repelled from this conductor through the observer's body to earth, leaving it positively charged. So long as *MN* remains close to *AB* the repulsion of the negative charge on the insulating base prevents electrons passing to *MN* from any other conductor which may be brought into contact with it. If, however, the metal disk is carried to a distance from *AB* by means of the insulating handle and the observer's finger is brought near to the edge *N*, the attraction of the positive charge on the disk for the electrons in the finger may be great enough to cause a small spark to pass to *MN*. Since the positive charge on the metal disk is held in position by the attraction of the negative charge on the wax base when the two are close together, the positive charge is said to be *bound* under these conditions. On the other hand, when the metal disk is removed to such a distance that the force due to the charge on the base becomes negligible, the charge on the disk becomes *free* and is shared with any conductor with which the disk is brought into contact. In the first position the metal disk may be laid on top of the base, for the two come into actual contact at only a few spots, and as the base is an insulator no appreciable part of the negative charge on its surface passes to the disk.

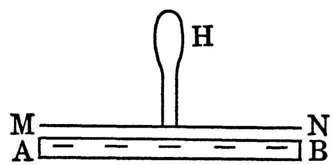


Fig. 4

By repeating the process of charging the metal disk of the electrophorus by induction and then transporting it to a distant

conductor a large charge may be imparted to the latter. Various machines have been devised to do this automatically. The *Wimshurst machine*, which is illustrated diagrammatically in Fig. 5, consists of two sets of metal carriers $a_1, a_2, \cdots$ and $b_1, b_2, \cdots$ which are revolved in opposite senses about an axis through O perpendicular to the plane of the figure. The stationary wire connectors AA and BB, provided with brushes to make contact with the carriers, may be considered to be connected to earth, although it is not necessary that they should be. Suppose that the carrier a_4 has been given a positive charge by contact with an electrified glass rod. It induces a negative charge on b_2 which is carried around to b_4 and there induces a positive charge on the carrier then at a_2. The carrier at b_4 then moves on to b_5, where it shares its charge with the collector C_1 which is provided with points to facilitate the transfer. The action of the points will be discussed in article 11. From C_1 the charge travels along the wire represented by the broken line to the terminal T_1.

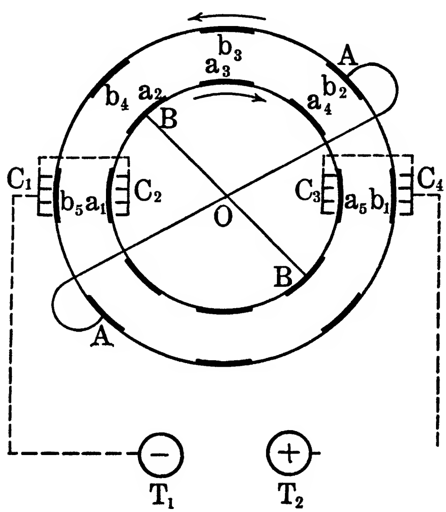


FIG. 5

In the meanwhile a_2 carries the positive charge which has been induced on it to a_4, where it induces a negative charge on the carrier then at b_2, and proceeding to a_5 shares its charge with the collector C_3 and the terminal T_2. On the lower halves of the circles the signs of the charges on the two sets of carriers are reversed. In this way a positive charge is built up on T_2 and a negative charge on T_1, and if the terminals are not too far apart a spark eventually passes across the gap between them. The

capacity of the terminals for accumulating electricity may be greatly increased by connecting them to the two plates of a device for storing charge known as a *condenser*. In this case discharges across the gap take place less frequently but with much greater intensity.

6. Coulomb's Law. — Before stating the law of attraction and repulsion between electric charges in quantitative form we must be careful to specify precisely the reference system or set of axes relative to which the quantities involved are to be measured. Consider a set of axes which is fixed relative to the average positions of the fixed stars. This set of axes, and in addition all sets of non-rotating axes which are in motion relative to it with constant velocities, are known as *inertial systems*. We shall suppose that the observer who is measuring the quantities involved in any of the laws of electromagnetism taken up in this book is permanently at rest in an inertial system to which we shall refer as the *observer's inertial system*. It is immaterial which one of the infinity of possible inertial systems we select in which to locate the observer. But in all cases the quantities to which we shall have occasion to refer are those measured relative to the particular set of axes chosen. The laws of electromagnetism which we are going to develop are valid relative to any inertial system in which the observer may happen to be located. When we pass from one inertial system to another the laws remain unchanged, although the values of the quantities involved may be quite different when determined by one observer from what they are when determined by another observer who is moving relative to the first. Thus one observer may conclude that there are electric forces in his vicinity but no magnetic forces, while another observer, located at the same spot as the first but moving relative to him, is aware of magnetic as well as of electric forces. The nature of an electromagnetic field — that is, a region in which electric and magnetic forces are present — is as much dependent upon the state of motion of the observer as upon the distribution of charged bodies and magnets. Lack of recognition of this

fact has led to much confusion in the solution of problems involving rotating magnets and the like.

In electrostatics we limit ourselves to the study of the forces between charges which are all at rest in the observer's inertial system. Suppose that we have two such charges q and q' at points located a distance r apart. If we double the charge q the force on q' evidently becomes twice as great, for each half of the doubled charge exerts the same force on q' as that exerted by the original charge. Therefore the force between the two charges, that is, the force exerted by either on the other, is proportional to q. In the same way we see that it is proportional to q'. Consequently *the force between two charges is proportional to the product of their magnitudes.*

To find how the force varies with the distance between two point charges Coulomb performed the following experiment in 1785. A charge is placed on a small conducting sphere B (Fig. 6) attached to the end of the insulating horizontal arm AB of a torsion balance. This arm is suspended at its center of mass from the drum D by means of the torsion fibre F. The torsion fibre is rigidly fastened to the drum at its upper end and to the horizontal arm at its lower end, so that if the drum is held fixed the arm cannot rotate about the fibre as axis without twisting it.

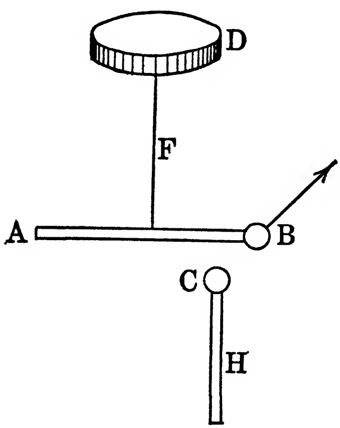


Fig. 6

A second charge is placed on the small conducting sphere C supported on the insulating handle H. We shall suppose this charge to be of the same sign as the charge already on B. If, now, C is brought up to a point at a predetermined distance in front of the original position of B, the charge on B is repelled and moves backward in the direction of the arrow, turning the arm AB and twisting the fibre F. Next the experimenter rotates the

drum in the opposite sense until B is restored to the position originally occupied. The angle through which it has been turned is equal to the twist in the fibre and therefore proportional to the force on B. By placing C at various distances in front of B Coulomb found that *the force between two charges is inversely proportional to the square of the distance between them.*

Analytically we can express Coulomb's law for the force F between two point charges by the formula

$$F = \frac{qq'}{r^2}. \tag{6–1}$$

If q and q' have the same sign F is positive, whereas if they have opposite signs F is negative. Consequently a positive value of the force indicates repulsion, a negative value attraction. As we shall see later the force between two charges is diminished by the presence of a dielectric between them. While the effect of air on the magnitude of the force is very small, the expression (6–1) applies exactly only to the case of charges *in vacuo.*

This formula contains implicitly the definition of the unit of charge. For if we make the charges equal, and make F and r each unity, $q = q'$ becomes unity. Therefore *the c.g.s. unit of charge is that charge which repels a like equal charge placed at a distance of 1 cm in vacuo with a force of 1 dyne.* Putting 1 gm cm/sec^2 for F and 1 cm for r and solving for $q = q'$ we find that the c.g.s. unit of charge is 1 gm$^{1/2}$ cm$^{3/2}$/sec. This unit, and all other c.g.s. units based on Coulomb's law of force, are known as *electrostatic units* and are designated by the abbreviation e.s.u.

Problem 6a. Two particles are suspended by strings of the same length l from the same point. Each has a mass m and a charge q. Show that the angle θ which each string makes with the vertical is given by

$$4mgl^2 \sin^3 \theta = q^2 \cos \theta.$$

7. Electric Intensity and Potential. — A region in which electric forces are acting is called an *electric field.* In order to explore the field we may carry a unit positive charge of very small dimensions around the field, being careful to keep all the charges

producing the field fixed in position so that the field shall not be changed by the attractions or repulsions exerted by the test charge. The force experienced by the unit test charge when at rest relative to the observer at any point in the field is known as the *electric intensity* E at that point. The unit of electric intensity is the *dyne per unit charge*. Since the electric intensity is a force per unit charge it is a vector and has direction as well as magnitude. In expressing vector relations we shall represent vectors by letters printed in **black face** type, using the same letters in *italics* to represent their scalar magnitudes. Thus the equation

$$\mathbf{C} = \mathbf{A} + \mathbf{B}$$

signifies that the vector $\mathbf{C}$ is the *vector sum* or *resultant* of the two vectors $\mathbf{A}$ and $\mathbf{B}$ as indicated in Fig. 7. The corresponding scalar relation is $C^2 = A^2 + 2AB \cos \alpha + B^2$, where α is the angle between the positive directions of A and B.

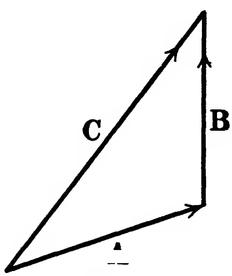


Fig. 7

Evidently the force F on a charge q placed at a point where the electric intensity is E is

$$F = qE. \tag{7-1}$$

If q is positive F has the direction of E, otherwise the opposite direction. It is clear from this equation that if the electric intensity is known at every point in the field the force on a charge of any magnitude placed at any point can be computed at once, provided the charges producing the field are held rigidly in position so that the field remains unaltered by the introduction of the charge on which the force is to be determined.

The electric intensity at a distance r from a point charge q is obtained at once from Coulomb's law by making q' equal to unity. It is

$$E = \frac{q}{r^2} \tag{7-2}$$

in the direction of the radius vector drawn from q. Since any field may be considered to be due to a number of point charges the resultant electric intensity at a point may be calculated by

finding the vector sum of the electric intensities due to the individual charges. Except in the simplest cases such a procedure is difficult to carry out, however, and simpler methods depending on the concept of potential are employed instead.

The *potential* V at a point in an electrostatic field is the work necessary to bring a unit positive charge from infinity, that is, from outside the field, up to the point in question, the charges producing the field being held rigidly in position during the process. Physically it is identical with the potential energy of a unit charge placed at the point in question and is measured in *ergs per unit charge.* As it represents work per unit charge potential is a scalar quantity.

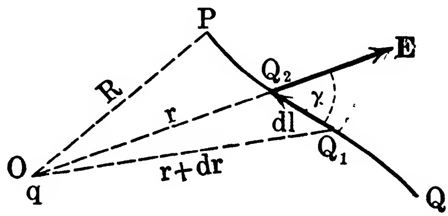


FIG. 8

Let us calculate the potential due to a point charge q placed at the origin. To find the potential at a point P (Fig. 8) distant R from q we must compute the work done in bringing a unit charge from infinity to P. Let QP be the path of approach to P. The electric intensity at Q_2 is

in the direction of the radius vector r and the work done against the repulsion of q in moving the unit charge the distance dl from Q_1 to Q_2 is

$$r^2 .$$

But

$$dl \cos \gamma = - dr$$

and consequently

$$dV = - \frac{q}{r^2} dr.$$

Integrating from ∞ to R,

$$V = - q \int_{\infty}^{R} \frac{dr}{r^2} = \frac{q}{R}, \tag{7-3}$$

showing that the value of the potential depends only upon the coordinates of the point P and is independent of the path followed in bringing the unit positive charge to P.

If the field is due to point charges $q_1, q_2, q_3, \cdots$ at distances $r_1, r_2, r_3, \cdots$ from P, the potential at P is the scalar sum

$$V = \frac{q_1}{r_1} + \frac{q_2}{r_2} + \frac{q_3}{r_3} + \cdots = \Sigma \frac{q}{r}, \tag{7-4}$$

and in the case of a continuous distribution of electricity consisting of ρ units of charge per unit volume occupying a volume τ and σ units of charge per unit area distributed over a surface s

$$V = \int_\tau \frac{\rho d\tau}{r} + \int_s \frac{\sigma ds}{r}. \tag{7-5}$$

As is evident from the forms of these expressions, *the electrostatic potential in all cases is a function only of the coordinates of the point P at which it is to be evaluated, and is independent of the path along which the unit charge is carried to P.* It is this property which makes the concept of potential important. In fact we say that a field of force possesses a potential only in regions where these conditions are fulfilled.

Evidently the difference in potential of two points in a field is measured by the work necessary to carry a unit positive charge from the one to the other, all paths between the same two end points being equivalent. The work, then, performed in carrying a charge q by any path from a point A where the potential is V_A to a point B where the potential is V_B is $q(V_B - V_A)$. If the charge is carried around a closed curve, the two ends of the path coincide. Therefore $V_B = V_A$ and the net work done is zero.

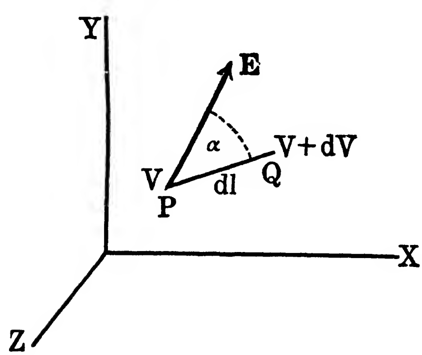


FIG. 9

Consider two nearby points P and Q (Fig. 9) a distance dl apart and denote the potentials at P and Q by V and $V + dV$ respectively. As the force on a unit positive charge is the electric intensity E, the *excess* of the potential at Q over that at P is the work done *against* the electric intensity in passing from P to Q, or the *drop* in potential in going from P to Q is the work done *by* the electric intensity. Therefore

$$dV = -E \cos \alpha dl.$$

Hence the component of E in the direction of the displacement PQ is

$$E \cos \alpha = -\frac{\partial V}{\partial l}, \tag{7–6}$$

equal to the space rate of decrease of potential in the direction PQ. The left-hand side of this equation is greatest for α equal to zero. Therefore the potential decreases most rapidly in the direction of the electric intensity.

If dl is parallel to the X axis it becomes dx and $E \cos \alpha$ becomes E_x, the X component of the electric intensity. Consequently

$$E_x = -\frac{\partial V}{\partial x}.$$

Similar expressions hold for the Y and Z components of E. So if V is expressed as a function of the coordinates x, y, z,

$$E_x = -\frac{\partial V}{\partial x}, \qquad E_y = -\frac{\partial V}{\partial y}, \qquad E_z = -\frac{\partial V}{\partial z}. \tag{7–7}$$

By means of these relations we can obtain the electric intensity if the potential function is known. As the electric intensity is equal to the space rate of decrease of potential, it follows that a positive charge placed in an electric field experiences a force urging it from regions of higher to regions of lower potential while a negative charge tends to move from regions of lower to regions of higher potential.

If the potential V is expressed as a function of the spherical coordinates r, θ, ϕ, where r represents the radius vector, θ the polar angle and ϕ the azimuth, the space rates of decrease of potential and therefore the components of electric intensity in the directions of increasing r, θ, ϕ are respectively

$$E_r = -\frac{\partial V}{\partial r}, \qquad E_\theta = -\frac{\partial V}{r\partial \theta}, \qquad E_\phi = -\frac{\partial V}{r \sin\theta\partial\phi}. \qquad (7\text{–}8)$$

Problem 7a. Charges of 200 and $-$ 100 e.s.u. are placed at the points (0, 0) and (1 cm, 0) in the XY plane. Find the components of electric intensity at the points (10 cm, 0) and (0, 10 cm) and find a point on the X axis where the field vanishes. Ans. $E_x = 0.765$ e.s.u., $E_y = 0$; $E_x = 0.098$ e.s.u., $E_y = 1.015$ e.s.u.; $x = 3.41$ cm.

Problem 7b. Find the potential due to the charges in the preceding problem at any point in the XY plane and obtain the rectangular components of the electric intensity by differentiation, checking the numerical values obtained above. What is characteristic of the potential function at the point of equilibrium found above? Is this a point of stable or unstable equilibrium?

Ans. $V = \dfrac{200}{\sqrt{x^2+y^2}} - \dfrac{100}{\sqrt{(x-1)^2+y^2}}$, saddle point, unstable.

Problem 7c. A thin glass rod of length l placed along the X axis with one end at the origin is electrified uniformly along its length with a total charge Q. Find the potential and the electric intensity at any point on the X axis beyond the end of the rod.

$$\text{Ans. } \frac{Q}{l}\log\frac{x}{x-l}, \quad \frac{Q}{x(x-l)}.$$

8. Gauss' Law. — Consider an element of surface ds at a point in an electric field at which the electric intensity is E. If α is the angle between the direction of E and that of the normal to the surface, the product $E \cos \alpha ds$ of the component of E perpendicular to the surface by the area of the surface is called the *electric flux* through the surface ds. The sign of the electric flux depends upon which side of the surface is chosen as positive, that is to say, upon whether the normal is drawn perpendicular to the surface on the one side or the other. The following conventions are employed to determine the positive sense of the normal:

(*a*) If the surface is closed, the outward drawn normal is taken as positive.

(*b*) If the surface is open, the positive sense of the normal is related to the sense in which the periphery is described by the rule that a right-handed screw with its axis perpendicular to the surface advances in the direction of the normal when rotated in the sense in which the periphery is described. This rule determines the positive sense of the normal only if the positive sense of describing the periphery is specified, or, *vice versa*, it prescribes the positive sense in which the periphery is to be described if the positive side of the surface is given.

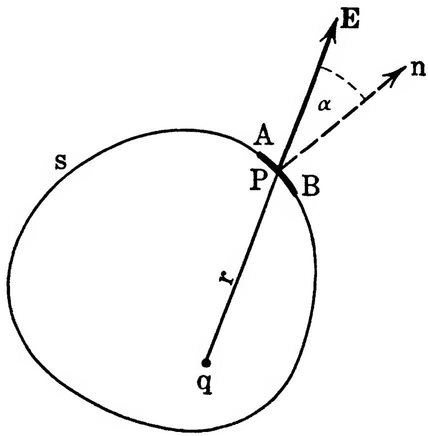


Fig. 10

Consider a point charge q (Fig. 10) inside a closed surface s. If n represents the normal to the surface at P, the electric flux dN through the element AB of area ds due to q is

$$dN = E \cos \alpha ds = q \frac{ds \cos \alpha}{\;}$$

If we draw straight lines from all points on the periphery of the surface element AB to q the cone so described is said to define a *conical angle* or *solid angle*. The solid angle is measured by the area intercepted on the surface of a sphere of unit radius having the vertex of the cone as center. Since the area of a sphere is proportional to the square of its radius, the magnitude of a solid angle is also equal to the area intercepted on the surface of any sphere with center at the vertex of the angle divided by the square of its radius. Furthermore, as the superficial area of a sphere

of radius r is $4\pi r^2$, the solid angle subtended at q by a surface such as s entirely surrounding it is 4π.

Now $ds \cos \alpha$ is the projection of the area AB perpendicular to the radius vector r, and the quotient of this projection by r^2 is the solid angle $d\Omega$ subtended by AB at q. Hence

$$dN = q\,d\Omega,$$

and if we sum up over the entire surface the total outward flux is seen to be

$$N = 4\pi q. \tag{8-1}$$

Consider now a charge q (Fig. 11) lying outside the closed surface. With q as vertex describe a cone of angular aperture

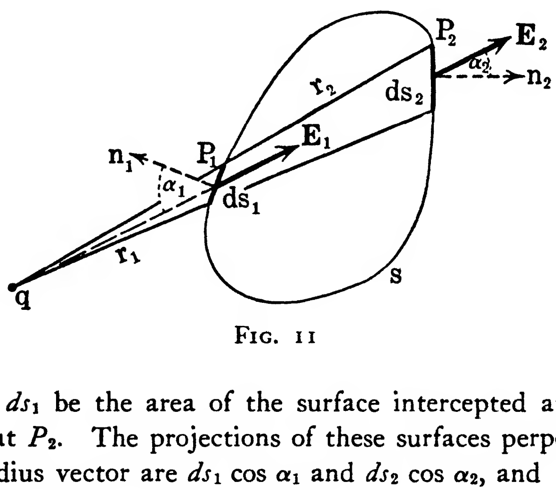


FIG. 11

$d\Omega$. Let ds_1 be the area of the surface intercepted at P_1 and ds_2 that at P_2. The projections of these surfaces perpendicular to the radius vector are $ds_1 \cos \alpha_1$ and $ds_2 \cos \alpha_2$, and

$$d\Omega = \frac{ds_1 \cos \alpha_1}{r_1^2} = \frac{ds_2 \cos \alpha_2}{r_2^2}.$$

Since the angle $\pi - \alpha_1$ between the directions of E_1 and n_1 is obtuse the flux through ds_1 is negative, signifying that it is directed inward through the closed surface instead of outward. Taking ds_1 and ds_2 together the outward flux is

$$-E_1 \cos \alpha_1 ds_1 + E_2 \cos \alpha_2 ds_2 = -q\,\frac{ds_1 \cos \alpha_1}{r_1^2} + q\,\frac{ds_2 \cos \alpha_2}{r_2^2}.$$

But the geometrical relation above shows that this expression vanishes. Since the whole surface s can be divided into pairs of elements subtending the same solid angle at q such that the inward flux through one annuls the outward flux through the other, the net outward flux through the entire surface due to a charge outside is zero. Therefore the flux through a closed surface is due to the charges enclosed by the surface alone.

If a number of point charges $q_1, q_2, q_3, \cdots$ are inside the surface s, the normal component of the resultant electric intensity is equal to the sum of the normal components of the electric intensities due to the individual charges and therefore the flux through the surface is

$$N = 4\pi(q_1 + q_2 + q_3 + \cdots) = 4\pi\Sigma q, \tag{8–2}$$

and if the charge is distributed continuously in the volume τ inside the surface s with density ρ units of charge per unit volume,

$$N = 4\pi \int_\tau \rho d\tau. \tag{8–3}$$

Equations (8–2) and (8–3) are equivalent statements of *Gauss' law*. In words this law states that the net outward electric flux through any closed surface is equal to 4π times the total charge contained within that surface. It is important to note that the deduction of this law depends only upon the fact that electrical forces vary inversely with the square of the distances between charges. Therefore it is valid for a gravitational field as well as for an electrostatic field.

If we write for N the surface integral of the normal component of the electric intensity in accord with the definition of electric flux, (8–2) and (8–3) take the forms

$$\int_s E \cos \alpha ds = 4\pi\Sigma q, \tag{8–4}$$

and

$$\int_s E \cos \alpha ds = 4\pi \int_\tau \rho d\tau. \tag{8–5}$$

In the first of these Σq represents the sum of all the charges contained in the closed surface s, and in the second the volume integral is to be taken over the entire volume τ surrounded by s.

9. Applications of Gauss' Law. — Gauss' law constitutes a powerful tool for finding the electric intensity in fields of such symmetry that we can draw surfaces everywhere normal to E at all points of which the magnitude of the electric intensity is the same.

Consider, for instance, a sphere s (Fig. 12) of radius a inside of which electric charge is so distributed that the charge density

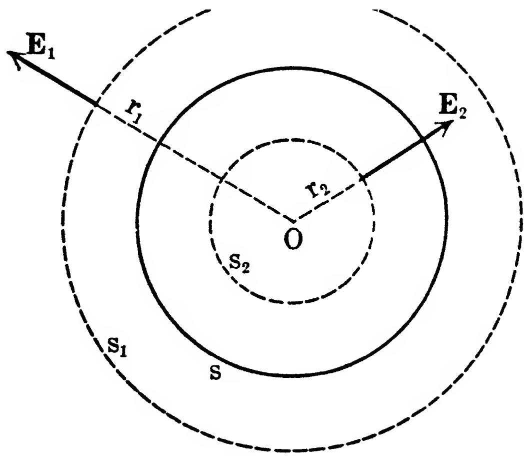


Fig. 12

ρ is a function of the radius vector r alone. Describe a concentric spherical surface s_1 of radius r_1 greater than a. It is clear from symmetry that the electric intensity is everywhere normal to s_1 and has the same magnitude E_1 at all points of this surface. Therefore the total flux through s_1 is $4\pi r_1^2 E_1$, and Gauss' law requires that

where Q represents the entire charge inside s, or

$$E_1 = \frac{Q}{r_1^2}. \qquad (9\text{–}1)$$

This is just the expression we should have found if all the charge had been concentrated at the center O of the sphere s. Therefore any distribution of charge in which the density is a function of the radius vector only produces the same field at exterior points as if the entire charge were located at its center.

If we apply Gauss' law to the spherical surface s_2 of radius r_2 less than a we find in the same way

$$E_2 = \frac{Q_2}{r_2^2},$$

where Q_2 is the portion of the charge inside s_2. In this case the electric intensity is that which would be produced by a charge Q_2 located at O. The charge lying between the spherical surfaces s_2 and s is without effect. If all the charge lay between these two surfaces there would be no field at points on the surface s_2 or at points inside this surface. Thus a charge spread uniformly over the surface of a sphere produces no field at points in its interior, although the field at exterior points is the same as if the entire charge were concentrated at the center of the sphere.

On the other hand, if the charge is distributed uniformly through the volume of the sphere,

$$Q_2 = \frac{r_2^3}{a^3} Q,$$

and

$$E_2 = \frac{Q}{a^3} r_2 \tag{9-2}$$

in the interior of the sphere.

Suppose we wish to find the force between two spherical charges Q_1 and Q_2 in each of which the charge density ρ is a function of the distance from the center only. Denote the centers of the two spheres by O_1 and O_2 and the distance between centers by R. Replace the second spherical distribution by a point charge Q_2 located at O_2. The electric intensity at O_2 due to the first sphere is

$$E_1 = \frac{Q_1}{R^2},$$

as proved above, and therefore the force on the point charge Q_2 at O_2 due to Q_1 is

$$F = \frac{Q_1 Q_2}{R^2} = Q_1 E_2,$$

where E_2 is the electric intensity at O_1 due to Q_2.

But the law of action and reaction requires that this expression should also represent the force exerted by Q_2 on Q_1. If now the point charge at O_2 is replaced by the spherical distribution originally assumed, the field due to Q_2 and therefore the force exerted by it on Q_1 remain unchanged. Consequently the force between the two spherical charges is the same as if each were a point charge located at its geometrical center. It is of interest to note that in order to deduce this result, which follows so simply from Gauss' law, for the corresponding case of gravitational attraction, Newton delayed publication of the law of gravitation for twenty years.

Next consider a uniformly charged plane MN (Fig. 13) of very great (strictly infinite) extent. Let σ be the charge per unit area of the surface. From symmetry it is clear that the electric intensity is perpendicular to the plane, being directed upward above the plane and downward below if the charge is positive. Describe a pill-box shaped surface $ABCD$, the flat bases AB and DC of the box lying parallel to and equidistant from the plane MN. If s is the area of one of the bases, the flux through the surface of the pill-box is $2Es$ and the charge enclosed is σs. Therefore Gauss' law takes the form

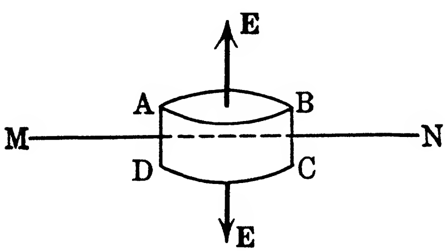


Fig. 13

$$2Es = 4\pi\sigma s,$$

or

$$E = 2\pi\sigma. \tag{9-3}$$

Consequently the field is uniform on each side of the plane, the magnitude of E being independent of the distance from the plane.

Finally consider the field due to two parallel conducting plates AB and CD (Fig. 14) of very great extent, the lower of which has a positive charge of density σ per unit area and the

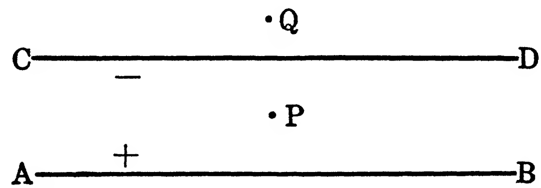


FIG. 14

upper of which has an equal negative charge. At a point P between the plates the electric intensity due to AB is $2\pi\sigma$ upward and that due to CD is also $2\pi\sigma$ upward. Therefore the total field is

$$E = 4\pi\sigma, \tag{9-4}$$

upward everywhere between the plates. The electric intensity at Q, however, is $2\pi\sigma$ upward due to AB and $2\pi\sigma$ downward due to CD. Hence the field at a point outside the plates, such as Q, vanishes.

Problem 9a. What is the potential due to a uniformly charged spherical shell of radius a at a point (a) outside, (b) on the surface, (c) in the interior of the shell? Ans. $\frac{Q}{r}, \frac{Q}{a}, \frac{Q}{a}$.

Problem 9b. The plane of Fig. 13 is replaced by a slab of thickness t, the charge being uniformly distributed throughout its volume. Find E at a distance y from the median plane of the slab less than one-half t. Ans. $4\pi\frac{\sigma}{t}y$.

Problem 9c. Find the field at a distance r greater than a from the axis of an infinitely long straight cylindrical rod of radius a which has a charge λ per unit length. Find the field inside the rod (a) if the charge is distributed uniformly over the surface of the cylinder, (b) if the charge is distributed uniformly throughout its volume. Ans. $\frac{2\lambda}{r}$, 0, $\frac{2\lambda}{a^2}r$.

Problem 9d. Prove that the electric intensity outside a uniformly charged spherical shell is the same as if the charge were concentrated at the center of the shell by finding the potential due to each element of charge and integrating over the surface of the shell. By the same method show that the field inside the shell vanishes.

10. Lines of Force. — A *line of force* in an electric field is a curve so drawn as to have everywhere the direction of the electric intensity, the sense in which the line is described being indicated on diagrams by an arrow-head. Free positive charges, therefore, tend to move along lines of force in the forward sense and free negative charges in the backward sense. A bundle of M lines of force, where M is a large integer arbitrarily chosen but fixed in value, is known as a *tube of force*. Lines of force are drawn in such density that the number of tubes of force per unit cross-section is everywhere equal to the magnitude of the electric intensity. By selecting a large enough number for M the representation of an electric field by lines of force may be made as nearly continuous as desired, even in regions where the electric intensity is very small.

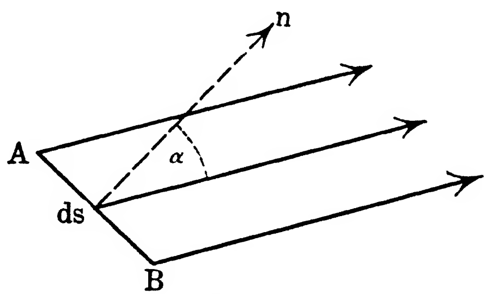


FIG. 15

Consider a small surface AB (Fig. 15) of area ds through which dN tubes of force pass at an angle α with the normal n. The cross-section of this group of tubes is the projection of ds on a plane perpendicular to E, that is, $ds \cos \alpha$. Therefore, in accordance with the convention for drawing lines of force,

$$E = \frac{dN}{ds \cos \alpha},$$

or

$$dN = E \cos \alpha ds.$$

But the right-hand side of this equation represents the electric flux through ds. Therefore *the number of tubes of force passing through a surface is equal to the electric flux through that surface.*

We shall now prove two fundamental theorems regarding lines of force.

Theorem I. Lines of force are continuous in regions containing no charge. To prove this theorem consider the closed surface AB (Fig. 16) consisting of a tube bounded by lines of force and

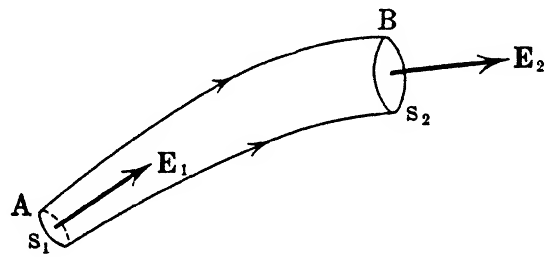


FIG. 16

terminated by the cross-sections s_1 and s_2. The flux through the sides of the tube vanishes, because the electric intensity is everywhere parallel to the surface. Therefore the total flux is that through the ends of the tube, and as there is no charge inside the tube Gauss' law requires that the inward flux through s_1 should equal the outward flux through s_2. But the flux through a surface has been shown to be equal to the number of tubes of force passing through the surface. Therefore as many tubes of force pass out through s_2 as come in through s_1. Consequently lines of force are continuous in the region occupied by the section of tube under consideration and hence in any region containing no charge.

Theorem II. The number of tubes of force diverging from a positive charge q or converging on a negative charge $-q$ is $4\pi q$. This theorem follows at once from Gauss' law. For the number of tubes of force passing through any closed surface surrounding

a charge q is equal to the electric flux through the surface, and the latter is equal to $4\pi q$. Each tube starts on a positive charge of magnitude $1/4\pi$ and ends on an equal negative charge.

In Fig. 17 the lines of force are shown for the fields due to (*a*) two equal point charges of the same sign, (*b*) two equal point charges of opposite sign, (*c*) two oppositely charged parallel conducting plates, (*d*) a point charge near an earthed conducting plate.

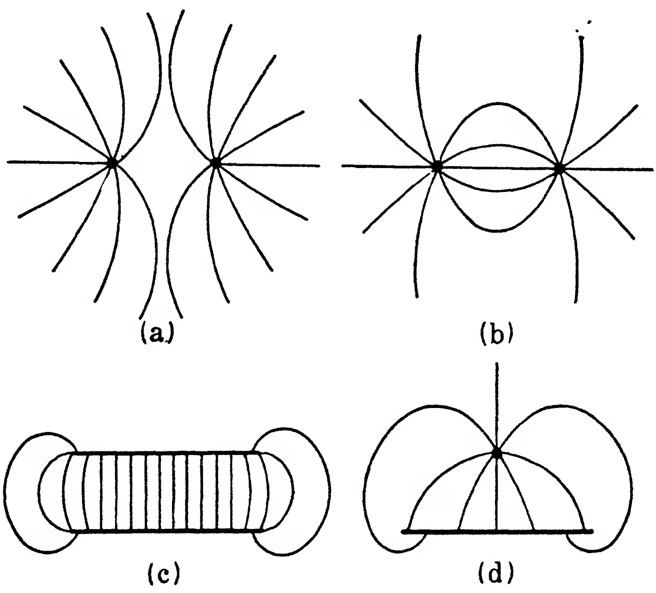


FIG. 17

An important theorem due to Earnshaw states that no charge can be in stable equilibrium in an electrostatic field under the influence of electrical forces alone. The proof of this theorem is very simple if we make use of the properties of lines of force which we have just developed. Assume that O (Fig. 18) is a point at which a positive test charge would find itself in stable equilibrium in the field under consideration. Then in whatever direction the test charge is displaced from O the force due to the field must urge it back toward O. Consequently the lines of

force must converge at O. But this would require the presence of a negative charge. Therefore there can be no point of stable equilibrium in an electrostatic field. A point of unstable equilibrium may exist, such as was found in problem 7*b*. In this case displacements of a positive charge along the X axis tend to increase whereas those in the plane at right angles give rise to restoring forces directed toward the point of equilibrium.

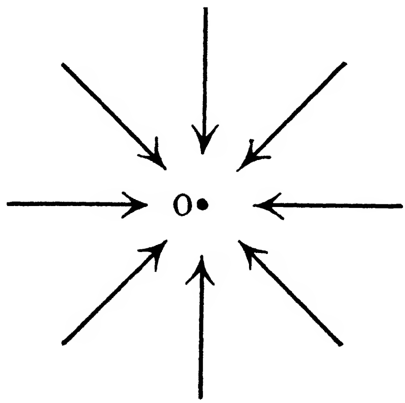


FIG. 18

11. Equipotential Surfaces.—An *equipotential surface* is a surface all points of which are at the same potential. No two equipotential surfaces can intersect, for if they did the potential along the curve of intersection would have more than one value. At points of equilibrium, however, a single equipotential surface intersects itself. Equipotential surfaces are drawn for equal increments of potential, for example, for the potentials 0, 5, 10, 15, 20, $\cdots$. As the electric intensity is equal to the space rate of decrease of potential, the field is most intense where the equipotential surfaces are most closely crowded together and least intense where they are farthest apart. Let AB and CD (Fig. 19) be two adjacent equipotential surfaces, AB being the surface of higher potential. If we pass from AB to CD along the normal PQ the space rate of decrease of potential is greater than if we follow any other path such as PR. Therefore the electric intensity, since it has the direction of the greatest space

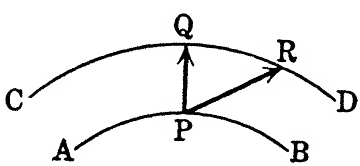


FIG. 19

rate of decrease of potential, is perpendicular to the equipotential surfaces and directed from surfaces of higher potential toward those of lower potential. Furthermore, lines of force, since they have everywhere the direction of the electric intensity, intersect equipotential surfaces orthogonally. They are most dense where the equipotential surfaces are closest together, for there the field is most intense.

Consider a charged conductor (Fig. 20) with an empty cavity inside. While the conductor is being charged the free electrons

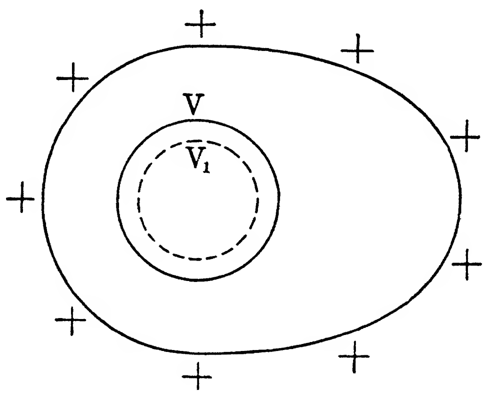


FIG. 20

in its interior move under the influence of electrical forces until they so distribute themselves that the electric intensity is zero everywhere inside the conductor. Therefore in the equilibrium state there is no electric field in the body of the conductor. Consequently there can be no charge in the interior of the conductor, for the presence of charge requires diverging or converging lines of force and therefore the existence of an electric field. So all the charge must be located on the surface of the conductor. Finally the absence of an electric field in the body of the conductor means that the potential is everywhere the same. The outer surface and the surface of the cavity are equipotential surfaces of the same potential.

Let us next investigate the field inside the cavity. Let V be

the potential of the conductor and V_1 that of the adjacent equipotential surface within the cavity. If V_1 is greater than V lines of force are directed from V_1 toward V at all points of the equipotential surface V_1. Therefore the flux through this closed surface is positive. But a positive flux requires a positive charge inside the surface V_1. Consequently V_1 cannot be greater than V. A similar line of reasoning shows that V_1 cannot be less than V. Therefore V_1 is equal to V, and the same is true of the potential at all points inside the cavity. Hence there is no field inside the cavity, and, as there is no field in the body of the conductor, it follows that there can be no charge on the surface of the cavity. The entire charge is located on the outer surface of the conductor.

If the force between charges did not vary inversely with the square of the distance the field inside a cavity in a charged conductor would not vanish. Cavendish and later Maxwell attempted to detect the presence of electrical forces in such a cavity with null results. So sensitive is this method of showing that the force varies inversely with the square of the distance that Maxwell calculated that if the exponent differed from 2 by as much as one part in 21,600 evidence of a field would have been found.

Next let us calculate the field just outside the conductor in terms of the charge σ per unit area on its surface. As lines of force are perpendicular to equipotential surfaces the electric intensity just outside a conductor must be at right angles to its surface. Therefore the flux through the pill-box shaped surface $ABCD$ (Fig. 21) enclosing a small area s of the surface consists entirely of that through the base AB, the flux through CD being zero since there is no field in the body of the conductor. Consequently Gauss' law gives

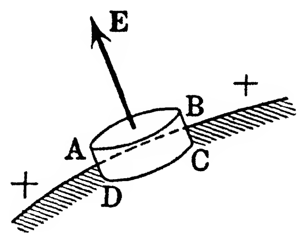


FIG. 21

$$Es = 4\pi\sigma s,$$

or

$$E = 4\pi\sigma. \qquad (11\text{–}1)$$

Finally we shall calculate the stress S on the surface of the conductor due to the action of the field outside on the charge σ located on a unit area of the surface. Strictly speaking the charge does not lie on a mathematical surface but occupies a thin layer extending a short distance into the conductor. Let us describe in this layer a number of surfaces parallel to the surface of the conductor so spaced that the charge $d\sigma'$ per unit area included between each pair of adjacent surfaces is the same. Then, since lines of force originate on charges, the electric intensity increases by a constant amount in passing from each surface to the next, starting from zero at the innermost surface. If σ' is the charge per unit area between a given surface and the innermost surface the electric intensity at the former is

$$E' = \frac{\sigma'}{\sigma} 4\pi\sigma = 4\pi\sigma'.$$

Therefore the tension stress on the surface of the conductor is

$$S = \int_0^\sigma E'd\sigma' = 4\pi \int_0^\sigma \sigma' d\sigma' = 2\pi\sigma^2. \qquad (11\text{–}2)$$

Since σ appears as a square in this expression the stress on the surface is always a tension, no matter whether the charge is positive or negative. In terms of the electric intensity $E = 4\pi\sigma$ just outside the surface of the conductor the tension is

$$S = \frac{1}{8\pi} E^2. \qquad (11\text{–}3)$$

Faraday and Maxwell conceived a tension of this amount as existing everywhere along the lines of force. On this representation the lines of force may be thought of as having the properties of stretched elastic bands tending to draw together the positive charges on which they originate and the negative charges on which they terminate. With this conception in mind mere inspection of the fields illustrated in Fig. 17 reveals the

direction in which each of the charges or conductors depicted there tends to move under the action of the electrical forces operating on it.

Let us examine qualitatively the field between two oppositely charged parallel conducting planes AB and CD (Fig. 22) the latter of which is provided with a rounded protuberance M and a point P. The lines of force are represented by full lines while the traces of the equipotential surfaces are shown by broken lines. While the field is somewhat increased in the neighborhood of M, it is seen to be very intense close to P. Therefore air molecules in the vicinity of the point are strongly attracted to it. On reaching P they acquire some of the charge on the point and are then repelled to AB. In this way the plate CD loses its charge to AB through the agency of the point. This action is made use of in transferring the charge from the carriers to the collectors of an induction machine such as was described in article 5.

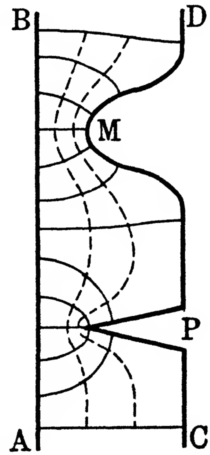


Fig. 22

Problem 11a. Construct roughly the lines of force and the equipotential surfaces in the field of the two charges of problem 7*b*.

Problem 11b. Two concentric spherical metal shells of radii 2 cm and 8 cm have charges of 100 e.s.u. and 5000 e.s.u. respectively. Find the stress on each. Ans. 24.9 dyne/cm^2, 253 dyne/cm^2.

CHAPTER II

DIELECTRICS AND CONDUCTORS

12. Electric Dipoles. — If two equal point charges of opposite sign are located at the same point, the field of the one will exactly annul that of the other so that the pair will give rise to no electrical forces on a third charge. If, however, one of the two charges is displaced a small distance relative to the other the fields of the two will no longer quite compensate. Such a combination is known as an *electric dipole* or an *electric doublet.* We have already made note of the fact that when a dielectric is subjected to electrical forces the bound electrons in each atom suffer a small displacement. Therefore each atom becomes an electric dipole, or more accurately a group of electric dipoles

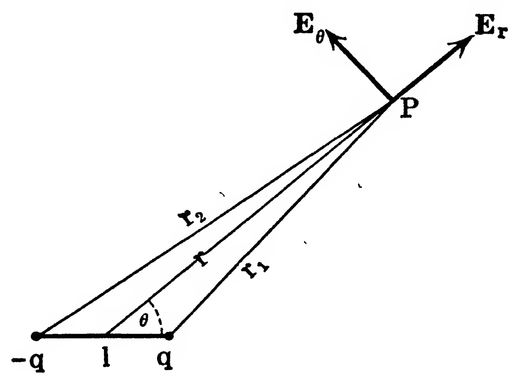


FIG. 23

each one of which may be supposed to consist of one electron and one proton. Under such circumstances the dielectric is said to be *polarized.*

We shall suppose that the distance l (Fig. 23) between the two charges q and $-q$ constituting a dipole is very small and shall calculate the potential and the components of electric

intensity at a point P at a distance r from the center of the dipole large compared with its length l. The potential at P due to the two charges q and $-q$ is

$$\frac{q}{r_1} - \frac{q}{r_2} = \frac{q}{r - \frac{l}{2}\cos\theta} - \frac{q}{r + \frac{l}{2}\cos\theta} = \frac{ql\cos\theta}{r^2 - \frac{l^2}{4}\cos^2\theta}$$

$$= \frac{ql\cos\theta}{r^2},$$

as l^2 is negligible compared with r^2. The product ql is known as the *electric moment* of the dipole and is designated by p. Evidently the electric moment is a vector having the direction of the axis of the dipole. We shall take its positive sense to be that of the line from $-q$ to q. Then the numerator of the expression for the potential is just the component of the electric moment in the direction of the radius vector r and we can write

$$V = \frac{p\cos\theta}{r^2}. \qquad (12\text{–}1)$$

If a number of dipoles are located in a small region the potential at a distant point P is given by the same expression provided

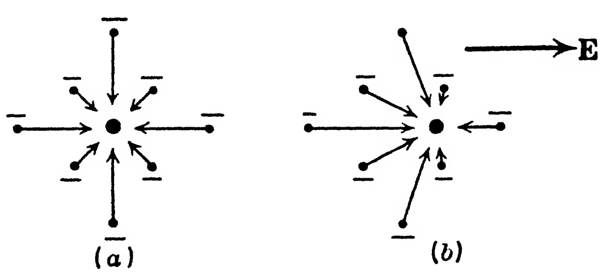


FIG. 24

we understand by p the magnitude of the vector sum of the electric moments of all the dipoles and by θ the angle which the resultant electric moment makes with the radius vector r. In the case of a normal atom (Fig. 24a) the electrons are distributed symmetrically about the protons in the nucleus, and if we pair

off each electron with a proton we have a number of dipoles the moments of which are represented by arrows in the figure. Clearly the resultant moment is zero and therefore the potential vanishes at distant points. If the atom is placed in an electric field, however, the electrons are displaced relative to the nucleus in a direction opposite to that of the electric intensity E (Fig. 24*b*). Evidently in this case the atom has a resultant electric moment which has the same direction as the field if the atom is isotropic.

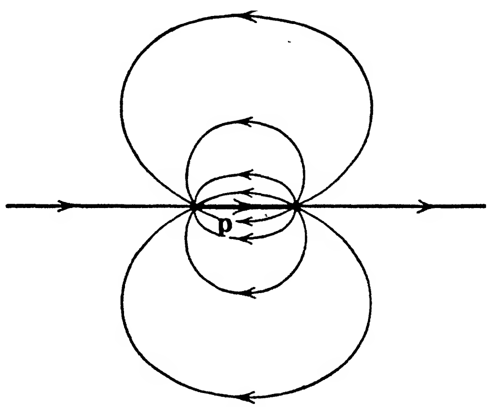


FIG. 25

Differentiating the expression (12–1) for the potential we find for the components E_r and E_θ of the electric intensity in the directions of increasing r and increasing θ respectively

$$\left.\begin{aligned} E_r &= -\frac{\partial V}{\partial r} = \frac{2p\cos\theta}{r^3}, \\ E_\theta &= -\frac{\partial V}{r\,\partial\theta} = \frac{p\sin\theta}{r^3}. \end{aligned}\right\} \qquad (12\text{–}2)$$

The lines of force in the field of a dipole are shown in Fig. 25, which differs from Fig. 17*b* only in scale.

13. Density of Charge in a Polarized Dielectric. — Although the electric field undoubtedly changes very greatly in both magnitude and direction as we pass across a single atom in a

dielectric, electrical instruments are unable to measure the fluctuations which occur in so short a distance. The measurements we make have to do rather with the mean values of the field and of the charge averaged over regions containing many thousands of atoms. We shall assume a chaotic arrangement of atoms in a dielectric and calculate the average density of charge in a small volume τ of dimensions Δx, Δy, Δz which, however, is large enough to contain a great many atoms. Evidently the mean density of charge vanishes if the medium is unpolarized, so our problem is to calculate the excess of the charge entering τ over that leaving this volume when the bound electrons in each atom become displaced by an impressed electric field.

As the effect we are investigating is due to the displacement of charge of one sign relative to that of the other, the analysis is simplified without in any way limiting the generality of the result by assuming that the negative electricity remains fixed and that the electric moment of each atom is due to a displacement R of the positive electricity. If $ABCD$ (Fig. 26) represents a face $\Delta y \Delta z$ of the volume τ the positive charge passing through this face when the medium is polarized is that contained inside the prism $ABCDEFGH$ of slant height equal to R. If R is not parallel to the X axis a portion of the charge passing through $ABCD$ near one edge, such as AB, may not remain in the volume τ, but this loss is compensated by the charge entering τ through the portion of the surface lying beyond the opposite edge. Denoting the X component of the displacement by R_x, the volume of the prism $ABCDEFGH$ is $R_x \Delta y \Delta z$ and the positive electricity

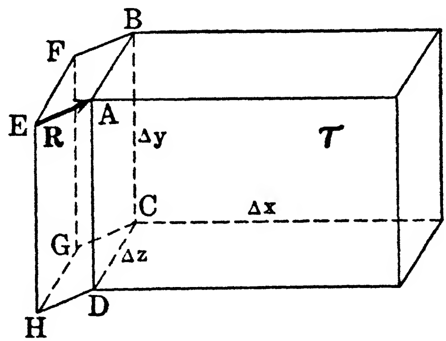


FIG. 26

contained in it is

$$\rho_1 R_x \Delta y \Delta z,$$

where ρ_1 represents the aggregate positive charge per unit volume. This expression, then, gives the charge passing into τ through the left-hand boundary when the medium is polarized. That passing out through the right-hand boundary is

$$\left\{ \rho_1 R_x + \frac{\partial}{\partial x}(\rho_1 R_x)\Delta x \right\} \Delta y \Delta z,$$

where ρ_1 as well as R_x may have different values at the two boundaries due to a change in the properties of the medium such as occurs in passing from one dielectric into another. Subtracting the second of these expressions from the first we find for the net charge entering τ through the two faces perpendicular to the X axis the quantity,

$$-\frac{\partial}{\partial x}(\rho_1 R_x)\Delta x \Delta y \Delta z.$$

Now $\rho_1 R$ is the electric moment per unit volume. This vector quantity is called the *polarization* and is denoted by P. So adding to the expression above corresponding expressions for the charge entering τ through the pairs of boundary surfaces perpendicular to the Y and Z axes, and dividing by the volume $\Delta x \Delta y \Delta z$ of the region τ, the net charge ρ_P per unit volume produced by the polarization of the medium is seen to be

$$\rho_P = -\left(\frac{\partial P_x}{\partial x} + \frac{\partial P_y}{\partial y} + \frac{\partial P_z}{\partial z} \right). \qquad (13\text{–}1)$$

If the polarization is constant throughout the medium the charge per unit volume in its interior is zero since each of the derivatives in (13–1) vanishes. In this case as much charge passes out of one face of a volume τ situated entirely inside the dielectric as passes in through the opposite one, and the only charges are those produced on the surface of the medium.

The surface charge is easily calculated in the case of a bar of length l and cross-section a polarized uniformly along its axis.

Denoting the charge per unit area on the two ends of the bar by σ_P and $-\sigma_P$ respectively the electric moment of the bar is $\sigma_P al$, and the electric moment per unit volume is σ_P. Therefore,

$$\sigma_P = P. \tag{13-2}$$

We can also express the charge passing into a volume τ when the dielectric is polarized as a surface integral. For if β (Fig. 27) is the angle between R and the outward drawn normal n to the surface s bounding τ, the charge entering τ through the surface element ds is that contained in the prism $ABCD$ of volume $-R\cos\beta ds$, that is, a charge $-\rho_1 R\cos\beta ds$. Therefore the total charge entering τ when the medium is polarized is

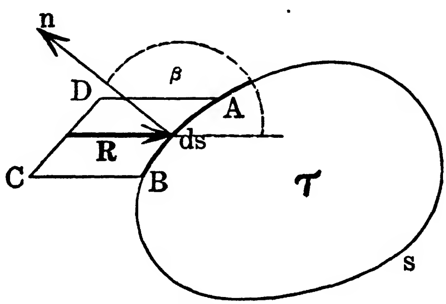


FIG. 27

$$q_P = -\int_s P\cos\beta ds, \tag{13-3}$$

where the surface integral is taken over the closed surface s which surrounds the volume τ.

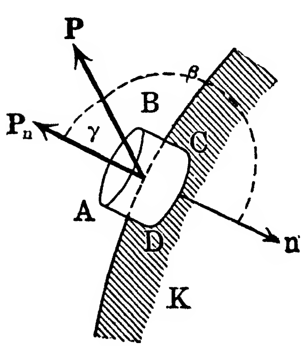


FIG. 28

We can use this formula to obtain a general expression for the density of charge on the surface of a dielectric. Let $ABCD$ (Fig. 28) be a pill-box shaped surface enclosing an area s of the surface of the dielectric K, the altitude BC of the pill-box being very small compared to the diameter of the base. Applying (13-3) to the surface of the pill-box, the value of the integral over AB vanishes since this base of the pill-box lies outside the dielectric where the polarization is zero, and the area of the

cylindrical portion of the pill-box is so small that the integral over this portion of the surface is negligible. So the right-hand side of (13–3) reduces to the integral over the base CD of the pill-box which lies inside the dielectric. If, then, σ_P is the charge per unit area on the surface of the dielectric,

$$\sigma_P s = -P \cos \beta s = P \cos \gamma s,$$

since γ is the supplement of the angle β between P and the normal n to CD. But $P \cos \gamma$ is the component P_n of the polarization at right angles to the surface of the dielectric. Therefore

$$\sigma_P = P_n \tag{13–4}$$

is the charge per unit area on the surface. This expression evidently reduces to (13–2) for the special case considered there. If the polarization of the medium is not uniform the surface charge is accompanied by a volume distribution of charge in the interior of the dielectric given by (13–1). To find the mean electric field in the vicinity of a dielectric — whether outside or inside the medium — we must add to the impressed electric intensity due to outside charges that produced by the surface charge (13–4) on the surface of the dielectric and that produced by the volume charge (13–1) in the interior of the dielectric.

Problem 13a. A sphere of radius a is polarized along the radius vector so that $P = kr$. Find the volume and surface charge densities and show that the total charge is zero. Ans. $-3k$, ka.

Problem 13b. A thin rod of cross-section a and length l lies on the X axis with its nearer end at a distance b from the origin. The polarization is kx parallel to the X axis. Find the charge at each end and the volume density of charge in the interior of the rod, and show that the total charge is zero. Also find the potential at the origin due to the rod. Ans. $-kab$, $ka(b+l)$, $-k$, $-ka \log\left(1+\frac{l}{b}\right)$.

Problem 13c. The following cavities are cut in a polarized dielectric; (*a*) a narrow slit at right angles to the polarization, (*b*) a sphere, (*c*) a long needle-like cavity with its axis parallel to the polarization. Find the electric intensity at the center of each due to the charge on the surface of the cavity. Ans. (*a*) $4\pi P$, (*b*) $\frac{4}{3}\pi P$, (*c*) 0.

Problem 13d. A long cylindrical cavity is cut in a polarized dielectric at right angles to the polarization. Find the electric intensity at its center due to the charges on the surface of the cavity. Ans.

14. Gauss' Law for Charges in a Dielectric. — Consider a dielectric in which are immersed free charges $q_1, q_2, q_3, \cdots$. Describe any closed surface s which may lie partly or wholly within the dielectric and may also surround some of the immersed charges. Gauss' law states that the electric flux through this surface is equal to 4π times the sum of the charges enclosed. Denoting the charge inside s due to the polarization of the dielectric by q_P,

$$\int_s E \cos \alpha ds = 4\pi(q_P +$$

where α is the angle between E and the outward drawn normal to the surface.

Replacing q_P by the right-hand side of (13–3) and transposing this term to the left-hand side of the equation so as to combine the two surface integrals,

$$(E \cos \alpha + 4\pi P \cos \beta)ds = 4\pi\Sigma q, \qquad (14\text{–}1)$$

where β is the angle between P and the outward drawn normal to the surface.

Using vector notation, we define the *electric displacement* * $\mathbf{D}$ as the vector sum of $\mathbf{E}$ and $4\pi\mathbf{P}$, that is,

$$\mathbf{D} \equiv \mathbf{E} + 4\pi\mathbf{P}. \qquad (14\text{–}2)$$

It is clear from (14–1) that E and P and consequently D have the same physical dimensions. All three, therefore, may be expressed in the same units.

If, now, γ is the angle between D and the normal n to the surface s,

$$D \cos \gamma = E \cos \alpha + 4\pi P \cos \beta,$$

* As here defined the electric displacement is 4π times as great as the conventional value of this quantity in electrostatic units. The definition here given is preferred on account of its closer analogy with the magnetic induction **B**.

as is clear from Fig. 29. Therefore Gauss' law (14–1) becomes

$$\int_s D \cos \gamma ds = 4\pi\Sigma q, \tag{14-3}$$

or, for a continuous distribution of free charge of density ρ per unit volume,

$$\int_s D \cos \gamma ds = 4\pi \int_\tau \rho d\tau,$$

the volume integral being taken over the entire volume τ enclosed by the surface s.

It is only in an anisotropic dielectric, such as crystalline quartz, that **P** and **E** have different directions. In the case of isotropic media, such as glass or paraffin, **P** and **D** have the same direction as **E**. Then the three angles α, β and γ in Fig. 29 are equal and (14–2) may be replaced by the scalar equation $D = E + 4\pi P$.

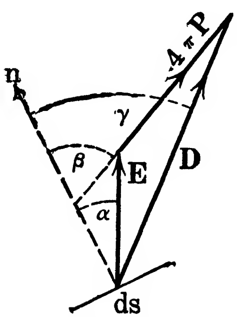


Fig. 29

We define the *flux of displacement* through a surface ds as the product of the component of the displacement D normal to the surface by the area of the surface. Then the left-hand side of (14–3) or (14–4) represents the outward flux of displacement through the closed surface s and these two equivalent forms of Gauss' law state that the total outward flux of displacement through any closed surface is equal to 4π times the sum of the *free charges* enclosed within the surface, where the term *free charges* designates all charges except those due to the polarization of the medium. The charges due to the polarization of the dielectric are included in the left-hand members of equations (14–3) and (14–4).

Lines of electric displacement may be drawn so as to have everywhere the direction of D and in such density that the number of *tubes of displacement* — defined as bundles of M lines of displacement — per unit cross-section is equal to the displacement. Following the same line of argument as that

applied to lines of force in article 10 it is clear that lines of displacement are continuous in regions in which no free charges are present, and that 4π tubes of displacement originate on each unit positive free charge and terminate on each unit negative free charge. When we pass from one dielectric to another through the surface separating them, or from a dielectric into empty space, lines of displacement are continuous provided no free charge resides on the surface. Lines of force, on the other hand, are not continuous across the boundary.

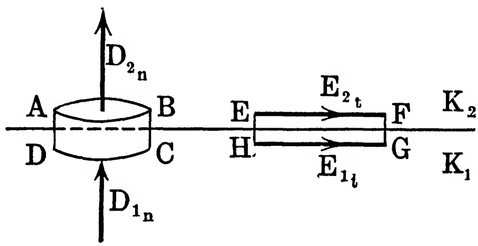


FIG. 30

Consider two dielectrics K_1 and K_2 (Fig. 30) in contact. Describe the pill-box shaped surface $ABCD$ about an area s of the surface of separation, the height BC of the pill-box being very small compared to the diameter of the bases. Let D_{1_n} and D_{2_n} be the normal components of the displacement in the two media. Then, as there is no free charge inside the pill-box, Gauss' law requires that

$$D_{2_n}s - D_{1_n}s = 0,$$

or

$$D_{2_n} = D_{1_n}. \tag{14-5}$$

Therefore the normal component of the displacement is the same on both sides of the surface of separation.

Next consider the rectangle $EFGH$ of length l and negligible height. Let E_{1_t} and E_{2_t} be the tangential components of the electric intensity in the two media. Then the work done in taking a unit charge around the rectangle is

$$E_{2_t}l - E_{1_t}l.$$

The principle of conservation of energy requires that this work should vanish, for otherwise we should have a perpetual motion mechanism. Hence

$$E_{2_t} = E_{1_t}, \tag{14-6}$$

and the tangential component of the electric intensity is the same on both sides of the surface.

15. Isotropic Dielectrics. — The results of the last article are equally valid for isotropic and for anisotropic media. Henceforth we shall confine our remarks to isotropic dielectrics. Since the polarization is produced by the electric field we should expect that P would be proportional in magnitude to E in such a medium as well as being in the same direction as E. If we write

$$P = \epsilon E, \tag{15-1}$$

the *electric susceptibility* ϵ is found to be practically constant for steady fields. Hence

$$D = E + 4\pi P = (1 + 4\pi\epsilon)E,$$

and if we put

$$\kappa \equiv 1 + 4\pi\epsilon, \tag{15-2}$$

the relation between D and E becomes

$$D = \kappa E. \tag{15-3}$$

The quantity κ is known as the *specific inductive capacity* of the dielectric, or the *dielectric constant.* Since D, E and P have the same physical dimensions ϵ and κ are pure numbers. Equation (15–3) tells us that there are κ times as many tubes of displacement per unit cross-section in an isotropic dielectric as there are tubes of force. Tubes of displacement terminate on *free charges* only, each tube starting on a free charge $1/4\pi$ and ending on a free charge $-1/4\pi$, whereas tubes of force terminate on both free and polarization charges.

In an isotropic dielectric Gauss' law (14–3) takes the form

$$\int_s \kappa E \cos \gamma ds = 4\pi\Sigma q, \tag{15-4}$$

and if the dielectric is homogeneous κ is a constant so that we may write

$$\int_s E \cos \gamma ds = \frac{4\pi}{\kappa} \Sigma q. \qquad (15\text{–}5)$$

Since q in these formulas represents free charge, we see from the second that if no free charge is present in the interior of a homogeneous isotropic dielectric the flux of E out of any closed surface lying wholly in the medium is zero. Consequently no polarization charge can be present in the interior of such a dielectric. All charges must reside on its surface.

By means of Gauss' law (15–5) we find at once the electric field in an isotropic dielectric surrounding a spherical charge Q in which the charge density is a function of the distance from the center only. The method is identical with that employed in article 9 in the case where no dielectric is present. Since Gauss' law (15–5) for charges immersed in a dielectric differs from the corresponding law (8–4) for charges in empty space only in the appearance of the factor κ in the denominator on the right, the expression (9–1) for the electric intensity in empty space must be replaced by

$$E = \frac{Q}{\kappa r^2} \qquad (15\text{–}6)$$

when the charge is immersed in a dielectric. This quantity represents the mean force per unit charge which would be experienced by a small charge present in the interstices between the atoms of the medium at a distance r from the center of Q.

The reason why the electric intensity is reduced in the ratio 1 to κ by the dielectric is made clearer if we examine the polarization charges in the medium. Eliminating ϵ from (15–1) and (15–2),

$$P = \frac{\kappa - 1}{4\pi} E, \qquad (15\text{–}7)$$

and making use of (15–6),

$$P = \frac{\kappa - 1}{\kappa} \frac{Q}{4\pi r^2}.$$

If a is the radius of the charge Q, the polarization charge per unit area of the cavity in the dielectric in which Q lies is

$$\sigma_P = -(P)_{r=a} = -\frac{\kappa - 1}{\kappa}\frac{Q}{4\pi a^2}$$

from (13–4), and the total polarization charge on the surface of the cavity is

$$Q_P = 4\pi a^2 \sigma_P = -\frac{\kappa - 1}{\kappa} Q.$$

The effective charge producing the field is the sum of Q and the polarization charge Q_P, that is,

$$Q - \frac{\kappa - 1}{\kappa} Q = \frac{Q}{\kappa}.$$

Consequently the electric intensity in the dielectric is only $1/\kappa$th of that which would be produced by Q alone.

On account of the factor κ in the denominator of (15–6) we must replace the expression (7–4) for the potential due to a number of point charges by

$$V = \sum \frac{q}{\kappa r} \tag{15–8}$$

when the charges are immersed in a dielectric. Similarly we have for the potential due to a continuous surface distribution of charge

$$V = \int_s \frac{\sigma ds}{\kappa r}. \tag{15–9}$$

Furthermore, if the charged surfaces discussed in articles 9 and 11 are surrounded by a dielectric, the expressions obtained for the electric intensity at outside points must be modified by the introduction of the factor κ into the denominator of each. Thus (11–1) for the electric intensity just outside a charged conducting surface becomes

$$E = \frac{4\pi\sigma}{\kappa}. \tag{15–10}$$

When two dielectrics are in contact both lines of force and lines of displacement are bent in passing from the one to the other. As was shown in the preceding article lines of displacement are continuous provided there is no free charge on the surface of separation. To calculate the law of refraction of these lines we make use of equations (14–5) and (14–6). Designating the angles which the lines of displacement make with the normal to the surface of separation of the two dielectrics of Fig. 30 by θ_1 and θ_2 we have

$$\kappa_2 E_2 \cos \theta_2 = \kappa_1 E_1 \cos \theta_1$$

from (14–5) and

$$E_2 \sin \theta_2 = E_1 \sin \theta_1$$

from (14–6). Dividing the one by the other,

$$\frac{\tan \theta_1}{\tan \theta_2} = \frac{\kappa_1}{\kappa_2}. \tag{15–11}$$

In passing from empty space into a dielectric, then, the lines of displacement are bent *away* from the normal to the surface,

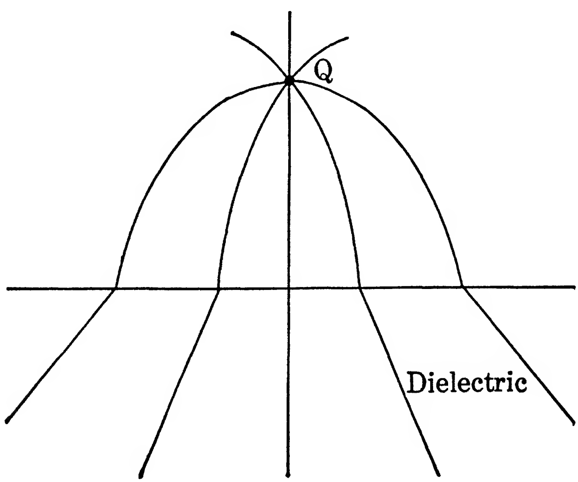


FIG. 31

an effect opposite to that which takes place when a ray of light passes into a transparent medium, such as glass. In Fig. 31

the bending of the lines of displacement is shown for the case of a point charge Q placed above a dielectric.

16. Stresses in a Dielectric. — In the last article we investigated the electric field produced by charges immersed in an isotropic dielectric. It remains to find the force experienced by a charge surrounded by a dielectric. As a first step we must examine the electrical forces on the elements of the dielectric itself.

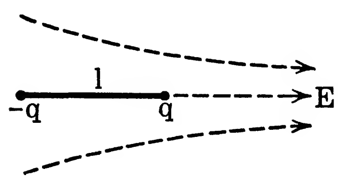


FIG. 32

As the polarization in an isotropic dielectric is parallel to the electric intensity each atomic dipole may be considered to be lined up with its axis parallel to the lines of force as in Fig. 32. If, then, E is the electric intensity at $-q$ that at q is $E + \frac{\partial E}{\partial l}l$ and the resultant force on the dipole is

$$-qE + q\left(E + \frac{\partial E}{\partial l}l\right) = ql\frac{\partial E}{\partial l},$$

where ql is the electric moment of the dipole. Replacing ql by the polarization P the force per unit volume is seen to be

$$F = P\frac{\partial E}{\partial l},$$

or

$$F = \frac{\kappa - 1}{4\pi}E\frac{\partial E}{\partial l} \tag{16-1}$$

from (15-7). This may be written

$$F = \frac{\kappa - 1}{8\pi}\frac{\partial}{\partial l}(E^2). \tag{16-2}$$

If, now, we designate the pressure in the dielectric by p, we see that the pressure over the face B of the rectangular parallelopiped of unit cross-section and thickness dl pictured in Fig.

33 must be greater than that over the face A by an amount just sufficient to balance the electrical force (16–2), that is,

$$\frac{\partial p}{\partial l}\,dl = \frac{\kappa - 1}{8\pi}\frac{\partial}{\partial l}(E^2)\,dl.$$

Integrating,

$$p = \frac{\kappa - 1}{8\pi}E^2 = \frac{1}{8\pi}\{DE - E^2\} \tag{16–3}$$

except for a constant of integration which is of no significance since it represents merely a constant pressure which is everywhere the same and therefore gives rise to no tractive forces. Equation (16–3) shows that the pressure in a dielectric is proportional to the square of the electric intensity, being large near charges, where E is great, and small at great distances from charges, where the field is weak.

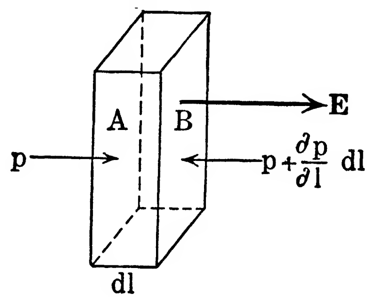


Fig. 33

Stress on Surface of Dielectric.—We have seen that a dipole well inside the body of a dielectric is urged by the electrical forces acting on it from the weaker to the stronger parts of the field. The same is true of the dipoles extending into the transition layer at the surface of a dielectric, where the polarization charge resides. However, the field changes so rapidly in this region from the value E in the interior of the dielectric to the value E_0 outside, that a special analysis of the forces is needed. We shall confine our discussion to the case where the field and therefore the polarization is normal to the surface of the dielectric.

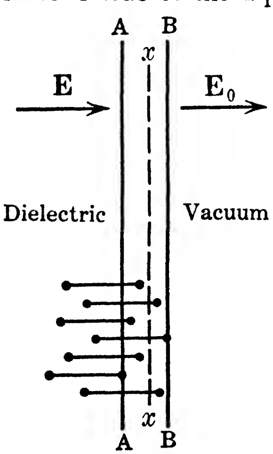


Fig. 34

Let the region between AA and BB (Fig. 34) be the transition layer at the surface of the dielectric lying to the left. As we pass from AA to BB the polarization charge increases from zero to its

full value σ_P per unit area, and since 4π lines of force originate on each unit of charge, the field increases from E to E_0 at the same rate that the polarization charge increases. Therefore, if the polarization charge per unit area between AA and any intermediate plane xx is σ_P', the field E' at xx is

$$E' = E + \frac{\sigma_P'}{\sigma_P}(E_0 - E). \tag{16-4}$$

Now the existence of the transition layer is due to the fact that the surface of the dielectric is not actually a perfectly smooth surface. At some places the last dipole reaches only as far to the right as AA, at others the last dipole in the dielectric extends well into the region between AA and BB, whereas at still others the last dipole reaches as far as BB. This is illustrated in the lower part of the figure. So the field acting on the right-hand or positive end of a surface dipole may be anything between E and E_0, depending upon how far the dipole under consideration extends into the transition layer. Since the polarization charge is due, however, to the uncompensated charges on the right-hand ends of the surface dipoles, the charge per unit area on the right-hand ends of all the surface dipoles which extend no farther to the right than xx is σ_P'. Hence the force on the right-hand ends of the dipoles with charge $d\sigma_P'$ is $E'd\sigma_P'$ and the force on the right-hand ends of all the surface dipoles is

$$\int_0^{\sigma_P} E'd\sigma_P' = E\int_0^{\sigma_P} d\sigma_P' + \frac{E_0 - E}{\sigma_P}\int_0^{\sigma_P} \sigma_P' d\sigma_P'$$
$$= \sigma_P E + \tfrac{1}{2}\sigma_P(E_0 - E)$$

per unit area. The left-hand or negative ends of the surface dipoles lie in the homogeneous portion of the dielectric where the field is E. So the force on them per unit area of the surface is $-\sigma_P E$. Adding this to the previous expression, we find for the force per unit area on the surface layer of dipoles

$$S_E = \tfrac{1}{2}\sigma_P(E_0 - E). \tag{16-5}$$

This, then, represents the electrical stress on the surface of the dielectric. If D is the electric displacement in the dielectric,

$E_0 = D = \kappa E$. Also $\sigma_P = P = (\kappa - 1)E/4\pi$. Hence we may write

$$S_E = \frac{(\kappa - 1)^2}{8\pi} E^2. \qquad (16\text{–}6)$$

To S_E we must add the mechanical pressure (16–3) in the homogeneous body of the dielectric to get the total stress S on the surface:

$$S = \frac{(\kappa - 1)^2}{8\pi} E^2 + \frac{\kappa - 1}{8\pi} E^2$$

$$= \kappa(\kappa - 1)\frac{E^2}{8\pi} = \frac{\kappa - 1}{\kappa}\frac{E_0{}^2}{8\pi}. \qquad (16\text{–}7)$$

As the expression for the stress involves only the square of the field strength it represents a positive stress or tension whether the electric intensity is directed along the outward normal to the surface, as we have supposed, or along the inward normal. This tension tends to stretch a solid dielectric in the direction of the lines of force of the field in which it is placed, an effect known as *electrostriction*. Using a liquid dielectric with an air bubble above it, Quincke has verified the formula (16–7) by measuring the increase in pressure of the air as the field is increased.

Stress on Surface of Conductor.—The stress on the surface of a conductor with charge density σ per unit area in empty space is $2\pi\sigma^2$ by (11–2). If the conductor is surrounded by a dielectric, the total tension S on its surface is the difference of this and (16–7), since the tension on the surface of the dielectric in contact with the conductor is transmitted as a pressure to the surface of the conductor.

Hence $$S = 2\pi\sigma^2 - \kappa(\kappa - 1)\frac{E^2}{8\pi}.$$

Now the free charge σ on the surface of the conductor may be expressed in terms of the electric intensity E or the electric displacement D in the dielectric just outside by (15–10), that is, $\sigma = \kappa E/4\pi = D/4\pi$. Hence

$$S = \frac{DE}{8\pi} = \frac{\kappa E^2}{8\pi} = \tfrac{1}{2}\sigma E = \frac{2\pi\sigma^2}{\kappa}. \qquad (16\text{–}8)$$

By integrating this stress over the entire surface of a conductor immersed in a homogeneous fluid dielectric we can find the resultant mechanical force experienced by the conductor.

Force on a Conductor.—Consider a conductor immersed in a homogeneous fluid dielectric. The field **E** at a point M (Fig. 35) in the dielectric just outside the surface element AB of the conductor may be considered as the resultant of three fields; the external field $\mathbf{E}_0$, the field $\mathbf{E}'$ due to the charge of density σ per unit area on AB and of density σ_P per unit area on the adjacent surface of the dielectric, and the field $\mathbf{E}''$ due to the charge on the remainder of the conductor and adjacent surface of the dielectric.

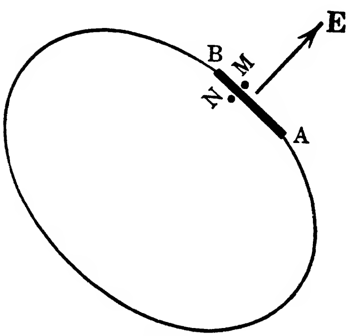


FIG. 35

Now the field at N due to the charge on the portion AB of the conductor and the adjacent surface of the dielectric is, from symmetry, equal and opposite to that at M, that is, equal to $-\mathbf{E}'$. The remainder of the field at N differs, however, only infinitesimally from that at M. But, as N is inside the conductor, the resultant field vanishes. Hence at N we have

$$\mathbf{E}_0 - \mathbf{E}' + \mathbf{E}'' = 0,$$

whereas at M we have

$$\mathbf{E}_0 + \mathbf{E}' + \mathbf{E}'' = \mathbf{E}.$$

Adding, we find that

$$\mathbf{E}_0 + \mathbf{E}'' = \tfrac{1}{2}\mathbf{E}. \qquad (16\text{–}9)$$

So the stress given by (16–8) is

$$\mathbf{S} = \tfrac{1}{2}\sigma\mathbf{E} = \sigma\mathbf{E}_0 + \sigma\mathbf{E}''.$$

The resultant force **F** on the conductor is obtained by integrating the stress **S** over its surface s, integration of a vector representing vector summation just as integration of a scalar represents scalar summation. Thus we obtain

$$\begin{aligned}\mathbf{F} &= \int_s \sigma\mathbf{E}_0 ds + \int_s \sigma\mathbf{E}'' ds \\ &= \int_s \sigma\mathbf{E}_0 ds + \kappa\int_s (\sigma + \sigma_P)\mathbf{E}'' ds\end{aligned}$$

since $\sigma + \sigma_P = \sigma/\kappa$. Now the second integral must vanish as it represents the sum of the forces exerted on each element of the conductor and adjacent dielectric by the charges on the remainder of the conductor and adjacent surface of the dielectric. Hence we have

$$\mathbf{F} = \int_s \sigma \mathbf{E}_0 ds. \qquad (16\text{–}10)$$

Therefore the resultant force exerted on a conductor immersed in a dielectric is obtained by multiplying each element of charge by the external electric intensity and summing up over the surface of the conductor. If, for instance, we have two conductors with charges Q_1 and Q_2 at a distance r apart large compared to their linear dimensions, the electric intensity at Q_2 due to Q_1 is

$$E = \frac{Q_1}{\kappa r^2},$$

and the resultant force on Q_2 is

$$F = \frac{Q_1 Q_2}{\kappa r^2}. \qquad (16\text{–}11)$$

As a charged conductor of very small dimensions is effectively the same thing as a point charge (16–11) also holds for the force between two point charges immersed in a dielectric.

Problem 16a. Two parallel plates of infinite extent are separated by a slab of solid dielectric of specific inductive capacity κ. One plate has a charge σ per unit area and the other a charge $-\sigma$. What force per unit area is required to move one plate away from the other? What is the stress on the surface of the dielectric? Ans. $2\pi\sigma^2$, $2\pi\sigma^2 \dfrac{\kappa - 1}{\kappa}$.

17. Electric Field in a Dielectric.—The electric field E_1 which is responsible for the displacement of the electrons of a single atom in a dielectric is the resultant of the field due to

external sources, including the surface and volume charges in the dielectric calculated in article 13, and that due to the surrounding polarized atoms. This is not, however, the total field E. The latter is the sum of E_1 and the field E_2 to which the polarized atom itself gives rise. In order to find the relation between E_1 and E we shall calculate E_2 and make use of the equation

$$E = E_1 + E_2. \tag{17-1}$$

Since the total field E is that due to external causes plus the surface and volume charges in the dielectric, E_2 may also be interpreted as the negative of the field exerted on the atom under consideration by the surrounding polarized atoms.

Since we are interested only in the mean value of the field we must calculate the value of E_2 averaged over the portion of the space occupied by the dielectric which pertains to a single atom. If there are n atoms per unit volume, the space belonging to each atom has a volume $1/n$. Furthermore, if the dielectric is isotropic, we can consider that the space belonging to a representative atom is spherical in form. Therefore we have to calculate the mean value of the electric intensity due to a dipole averaged over a sphere of volume $1/n$ inside of which the dipole is situated.

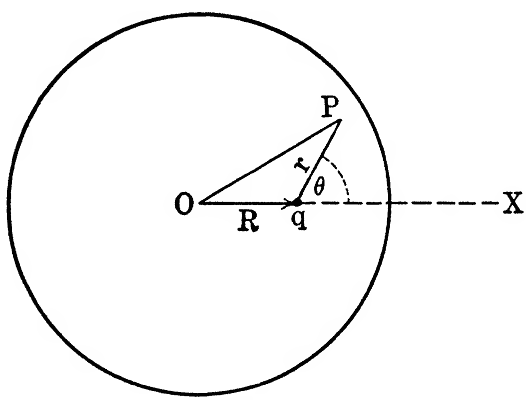


FIG. 36

Let O (Fig. 36) be the center of the sphere of volume $1/n$

which we have under consideration. We shall calculate the average value of the field due to a charge q which is displaced a distance R from O in the X direction. As it is clear from symmetry that the average values of the Y and Z components of the field vanish, it is necessary to calculate the average value of the X component only. At the point P the X component of the electric intensity due to q is

$$\frac{q}{r^2}\cos\theta.$$

To get the average value we must integrate this over the sphere and divide by the volume of the sphere. Hence

$$\begin{aligned} E_2 &= n\int\int \frac{q}{r^2}\cos\theta\cdot 2\pi r^2\sin\theta d\theta dr \\ &= 2\pi nq\int\int \cos\theta\sin\theta d\theta dr. \end{aligned}$$

The limits of θ are 0 and π, the lower limit of r is 0 and if a is the radius of the sphere the upper limit of r is given by

$$a^2 = R^2 + 2rR\cos\theta + r^2,$$

or

$$r = -R\cos\theta + \sqrt{R^2\cos^2\theta + a^2 - R^2},$$

where the positive sign has been taken before the radical since r is a positive quantity. Integrating with respect to r and putting in the limits,

$$E_2 = -2\pi nq\left\{R\int_0^\pi \cos^2\theta\sin\theta d\theta + \frac{1}{2}\int_0^\pi \sqrt{R^2\cos^2\theta + a^2 - R^2}\,d(\cos^2\theta)\right\}.$$

Since $\cos^2\theta$ has the same value at both limits the second integral vanishes and, in vector form,

$$\mathbf{E}_2 = -\frac{4}{3}\pi nq\mathbf{R}.$$

Now suppose we have inside the sphere two equal and opposite charges constituting a dipole. Let $\mathbf{R}_1$ be the vector dis-

tance of the positive charge q and $\mathbf{R}_2$ the vector distance of the negative charge $-q$ from O. Then the average field due to the dipole is

$$\mathbf{E}_2 = -\frac{4}{3}\pi n(q\mathbf{R}_1 - q\mathbf{R}_2) = -\frac{4}{3}\pi nq(\mathbf{R}_1 - \mathbf{R}_2).$$

But $q(\mathbf{R}_1 - \mathbf{R}_2)$ is the electric moment $\mathbf{p}$ of the dipole. So

$$\mathbf{E}_2 = -\frac{4}{3}\pi n\mathbf{p}.$$

As n represents the number of atoms per unit volume $n\mathbf{p}$ is the electric moment per unit volume, that is, the polarization $\mathbf{P}$. Consequently,

$$\mathbf{E}_2 = -\frac{4}{3}\pi\mathbf{P}. \tag{17-2}$$

Returning to (17–1) we have then

$$E_1 = E + \frac{4}{3}\pi P. \tag{17-3}$$

Since E_1 is the field responsible for the polarization of the atom under consideration we should expect the electric moment p of the atom to be proportional to E_1. Therefore we may write

$$p = \alpha E_1,$$

where the constant α is characteristic of the type of atom (or molecule) of which the dielectric is composed and independent, for instance, of the density of a gaseous dielectric. Therefore

$$P = np = n\alpha E_1. \tag{17-4}$$

Since the polarization P has the direction of E_1 equation (17–2) shows that the field E_2 to which a polarized atom itself gives rise is opposite to E_1. Therefore the polarization of an atom has the effect of weakening the field that would exist in its absence.

Now

$$E = \frac{4\pi}{\kappa - 1}P$$

from (15–7). Therefore (17–3) may be written

$$P = \frac{3}{4\pi}\frac{\kappa - 1}{\kappa + 2}E_1. \tag{17–5}$$

The coefficient of E_1 in this equation is the ratio of the polarization produced to the field actually operating on the polarized atoms. Eliminating P between (17–4) and (17–5) we get

$$\frac{\kappa - 1}{\kappa + 2} = \frac{4}{3}\pi n\alpha. \tag{17–6}$$

Since n represents the number of atoms per unit volume the right-hand side of this equation is proportional to the density of the dielectric. Therefore the ratio of $\kappa - 1$ to $\kappa + 2$ is also proportional to the density. That this is the case has been verified for a number of gases. The constant α may be shown to be approximately equal to the cube of the radius of the atom. Therefore the radius of the atom may be computed from measurements of the dielectric constant. The values so obtained are in fair agreement with those calculated from kinetic theory considerations.

It is important to distinguish the mean field E_2 produced by a dipole from the mean field which is produced by a single point charge. The latter vanishes on account of the symmetry of the radial field to which a point charge gives rise, and consequently we are justified in identifying the average electric intensity to which a point charge in the medium is subject with the total mean electric intensity E as we did in article 16 and shall do again in article 21.

18. Capacity and Condensers. — There is a definite relation between charge and potential in any electrostatic system. Consider first a single conductor with charge Q at potential V, surrounded by an infinite dielectric of specific inductive capacity κ. The interior of the conductor is an equipotential region, in which E is everywhere zero; so the absence of dielectric there is of no consequence and we may use (15–9) to calculate the poten-

tial at any point in space, including in particular, any point P on the surface of the conductor. Thus,

$$V = \int_s \frac{\sigma ds}{\kappa r}.$$

Evidently multiplying every element of charge by any factor multiplies V by the same factor. The surface of the conductor remains equipotential and equilibrium is not disturbed. That is, V is proportional to the total charge Q. If we write

$$Q = CV, \tag{18-1}$$

C is called the *capacity* of the conductor. It is the charge that can be placed on the conductor per unit potential.

In the case of a uniformly charged sphere of radius a immersed in an infinite dielectric the electric intensity outside the sphere is given by (15-6), that is,

$$E = \frac{Q}{\kappa r^2},$$

and the potential at the surface of the sphere is

$$V = -\int_\infty^a \frac{Q}{\kappa r^2} dr = \frac{Q}{\kappa a}. \tag{18-2}$$

Therefore from (18-1)

$$C = \kappa a. \tag{18-3}$$

As κ is a pure number this equation shows us that the electrostatic unit of capacity is the *centimeter*.

If several conductors are present there is no definite capacity associated with each one, since the charge on each depends not only on its potential, but also on the potentials of the other conductors. However, in the special case of two conductors arranged always to have charges of equal magnitude and opposite sign, we have as before

$$V_1 = \frac{Q}{C_1}, \qquad V_2 = -\frac{Q}{C_2},$$

so that we may write

$$Q = C(V_1 - V_2), \tag{18-4}$$

where

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2}.$$

Such a combination of conductors is called a *condenser.* It is a device for the storage of charge, the quantity stored being proportional to the potential difference of its elements. The constant C is known as the *capacity* of the condenser.

Concentric spheres with a dielectric between them (Fig. 37) form a simple condenser. Let the outer sphere be grounded, that is, at the same potential as the earth, walls of the room, and other near-by objects. Since we are concerned with potential difference alone, it is usually convenient to take the invariable potential of the earth as zero. As the charge $-Q$ on the outer sphere does not produce an electric field inside,

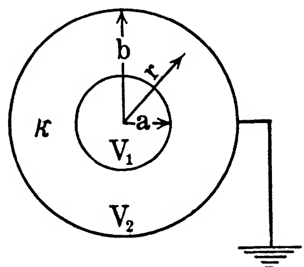


FIG. 37

$$V_1 - V_2 = -\int_b^a \frac{Q}{\kappa r^2}\,dr = \frac{Q(b-a)}{\kappa ab},$$

where a is the radius of the inner sphere and b the radius of the inner surface of the outer sphere. Therefore

$$C = \frac{\kappa ab}{b-a}, \tag{18–5}$$

which may be made much greater than the capacity (18–3) of an isolated sphere.

In the case just considered the charge on the outer sphere is evidently all on the inner surface. If now we ground the inner sphere instead of the outer, keeping $V_1 - V_2$ unchanged, there will be an additional charge on the outer surface of the outer sphere due to its capacity relative to surrounding objects.

This charge corresponds, approximately at least, to that on a single isolated sphere. So if b' is the radius of the outer surface, the total charge on the outer sphere is

$$-\frac{\kappa ab}{b-a}(V_1 - V_2) - b'(V_1 - V_2),$$

and the capacity is

$$\frac{\kappa ab}{b-a} + b'. \tag{18-6}$$

It is usually preferable to ground the outer sphere, since then the capacity is entirely independent of external objects. We have here a simple case of *electrostatic shielding*, to be discussed in more detail in article 19.

A very common and convenient type of condenser consists of two parallel plates. If the distance d between the plates is small compared to the dimensions of the plates, the electric intensity is zero outside and constant between the plates, except near the edges. In fact, in the central portion of the condenser E is perpendicular to the plates, and modifying (9–4) to take account of the dielectric,

$$E = \frac{4\pi\sigma}{\kappa}, \tag{18-7}$$

where σ is the charge per unit area. Then

$$V = Ed = \frac{4\pi\sigma d}{\kappa},$$

and the capacity per unit area is given by

$$C_1 = \frac{\kappa}{4\pi d}. \tag{18-8}$$

With the geometrical limitation mentioned above the *edge effect* is small. That is, if A is the area of either plate, the total charge is approximately $A\sigma$ and the total capacity is given by

$$C = \frac{\kappa A}{4\pi d}. \tag{18-9}$$

To obtain large capacity without excessive size condensers are often composed of a pile of plates separated by thin layers of dielectric such as mica or waxed paper. Alternate plates have the same polarity, so that double use is made of their area. If air is the dielectric, one group of plates may be arranged to move relative to the other, thus making the condenser *variable.*

When it is necessary to calculate capacity very accurately, in order to establish laboratory standards, or to make absolute measurements, the edge effect may be eliminated. Suppose the central portion of one plate of a single parallel plate condenser

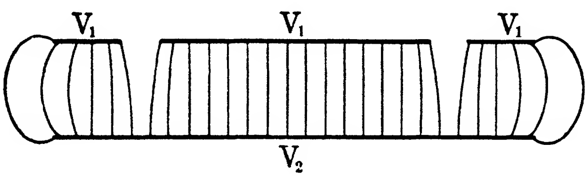


FIG. 38

is separated from the rest of the plate by a narrow gap (Fig. 38). The outer portion is called a *guard ring* and its use is due to Kelvin. The lines of force remain perpendicular almost to the edge of the guarded plate, and on the lower, unguarded plate they are practically undisturbed by the gap. If then A' equals the area A of the guarded plate plus one-half the area of the gap,

$$C = \frac{\kappa A'}{4\pi d} \tag{18–10}$$

to a high degree of accuracy. Of course, the potential of the ring must be kept exactly equal to that of the plate by some agency independent of that supplying the charge to the plate.

The capacity of combinations of condensers is easily calculated. Suppose we have several capacities C_a, C_b, $\cdots$ C_k in *parallel* (Fig. 39*a*). The total charge is given by

$$Q = Q_a + Q_b + \cdots Q_k = (C_a + C_b + \cdots C_k)(V_1 - V_2),$$

so that the resultant capacity is

$$C = C_a + C_b + \cdots C_k = \sum_a^k C_i. \qquad (18\text{–}11)$$

On the other hand with capacities in *series* (Fig. 39*b*),

$$V_1 - V_2 = V_a + V_b + \cdots V_k = \left(\frac{1}{C_a} + \frac{1}{C_b} + \cdots \frac{1}{C_k}\right) Q,$$

where V_a is the potential difference of the first pair of plates and so on. Therefore

$$\frac{1}{C} = \frac{1}{C_a} + \frac{1}{C_b} + \cdots \frac{1}{C_k} = \sum_a^k \frac{1}{C_i}. \qquad (18\text{–}12)$$

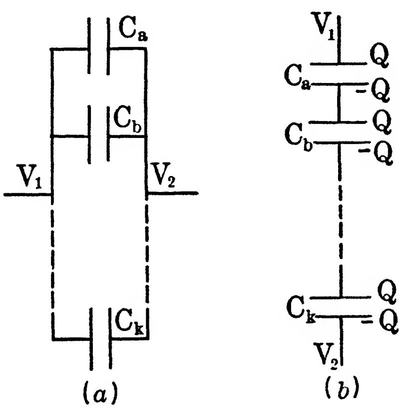


Fig. 39

It should be noted that when condensers are connected in series we usually have a combination of series and parallel capacities on account of the capacities of the outsides of the condensers to one another and to surrounding objects. Generally, however, these parallel capacities are negligible.

Problem 18a. Show that the capacity per unit length of two infinitely long concentric cylinders, of radii a and b ($b > a$), is $\dfrac{\kappa}{2 \log_e \dfrac{b}{a}}$, when the outer cylinder is grounded.

Problem 18b. A condenser known as a *Leyden jar* is formed by coating the inside and outside of a cylindrical bottle (open at the top) with tin foil. The diameter is 10 cm, the height 15 cm, and the thickness 1 mm. If the dielectric constant of the glass is 7, find the capacity. Ans. 3061 cm.

Problem 18c. A parallel plate condenser has a slab of dielectric whose thickness t is less than the separation of the plates. If the

latter quantity is d, find the capacity per unit area.

$$\text{Ans.}\quad \frac{1}{4\pi\left\{d - t\left(\frac{\kappa - 1}{\kappa}\right)\right\}}\,.$$

Problem 18d. A condenser is composed of a pile of n plates of alternate polarity, each of area A, separated by thin layers of dielectric of thickness d and dielectric constant κ. Find its capacity.

$$\text{Ans.}\quad \frac{(n-1)\kappa A}{4\pi d}\,.$$

Problem 18e. Given two identical spherical air condensers like that shown in Fig. 37. Let the outer sphere of one be connected to the inner sphere of the other, and the outer sphere of the latter connected to ground. Assuming that the spheres are not close to each other or to any other conductors, calculate approximately the total capacity C of the series combination, taking into account all stray capacity to ground.

$$\text{Ans.}\quad \frac{1}{C} = \frac{1}{\dfrac{ab}{b-a} + b'} + \frac{1}{\dfrac{ab}{b-a}}\,.$$

19. Coefficients of Potential, Capacity and Induction. — In the case of several charged conductors the relations between charges and potentials are still linear. For simplicity we shall confine the proof to three conductors, although the analysis is applicable to any number. Let V_1, V_2, V_3 be the potentials calculated from (15–9) for conductors (1), (2), (3) respectively when we put charges Q_1, Q_2, Q_3 on them, and V_1', V_2', V_3' the potentials when in place of the previous charges we put charges Q_1', Q_2', Q_3' on the three conductors. Then it follows from (15–9) that if we put charges $Q_1 + Q_1'$ on conductor (1), $Q_2 + Q_2'$ on (2) and $Q_3 + Q_3'$ on (3) the respective potentials are $V_1 + V_1'$, $V_2 + V_2'$ and $V_3 + V_3'$, for we have now just a superposition of the two separate fields.

In the special case that the primed Q's are equal to the unprimed Q's the primed V's are equal to the unprimed V's, showing that the effect of doubling all the charges is to double

all the potentials. Hence if we increase the charges on the three conductors in the same ratio, the potentials are also increased in that ratio. Consider now a set of charges Q_1, 0, 0 on the three conductors. As the charges on (2) and (3) are zero, all three potentials must be proportional to Q_1, so that we may write for the potentials of the three conductors $p_{11}Q_1$, $p_{21}Q_1$, $p_{31}Q_1$ respectively, where the p's are constants independent of Q_1. Similarly a set of charges 0, Q_2, 0 results in potentials $p_{12}Q_2$, $p_{22}Q_2$, $p_{32}Q_2$ and a set 0, 0, Q_3 in potentials $p_{13}Q_3$, $p_{23}Q_3$, $p_{33}Q_3$. In view of the conclusions reached in the preceding paragraph, then, the potentials when all three conductors are charged are

$$\left.\begin{aligned} V_1 &= p_{11}Q_1 + p_{12}Q_2 + p_{13}Q_3, \\ V_2 &= p_{21}Q_1 + p_{22}Q_2 + p_{23}Q_3, \\ V_3 &= p_{31}Q_1 + p_{32}Q_2 + p_{33}Q_3. \end{aligned}\right\} \quad (19\text{–}1)$$

The p's are known as *coefficients of potential.* Their values depend on the dielectric constant and the geometry of the system. Note that since the Q's represent total algebraic charge, the equations give no information about the distribution of charge on a conductor.

The coefficients are not all independent Suppose that charges Q_1, Q_2, Q_3 give rise to potentials V_1, V_2, V_3, and that Q_1', Q_2', Q_3' give rise to V_1', V_2', V_3'. Then

$$Q_1'V_1 + Q_2'V_2 + Q_3'V_3 = \int_{s_1} \sigma' ds' \int_{s_{123}} \frac{\sigma ds}{\kappa r_1} + \int_{s_2} \sigma' ds' \int_{s_{123}} \frac{\sigma ds}{\kappa r_2} + \int_{s_3} \sigma' ds' \int_{s_{123}} \frac{\sigma ds}{\kappa r_3},$$

where the s-subscripts indicate the surfaces over which the integration is to be performed. And similarly

$$Q_1V_1' + Q_2V_2' + Q_3V_3' = \int_{s_1} \sigma ds \int_{s_{123}} \frac{\sigma' ds'}{\kappa r_1} + \int_{s_2} \sigma ds \int_{s_{123}} \frac{\sigma' ds'}{\kappa r_2} + \int_{s_3} \sigma ds \int_{s_{123}} \frac{\sigma' ds'}{\kappa r_3}.$$

Evidently both of these integral expressions equal

$$\int_{s_{123}}\int_{s_{123}} \frac{\sigma\sigma'\, ds\, ds'}{\kappa r},$$

and so

$$\Sigma Q'V = \Sigma QV', \tag{19–2}$$

a theorem due to Green. If now we take

$$\begin{aligned} Q_1 &= 1, & Q_2 &= 0, & Q_3 &= 0, \\ Q_1' &= 0, & Q_2' &= 1, & Q_3' &= 0, \end{aligned}$$

we have at once $V_1' = V_2$, and from (19–1) $p_{12} = p_{21}$. Similarly, in general,

$$p_{ij} = p_{ji}, \tag{19–3}$$

a result of some importance. In words it is: *The potential of one conductor due only to a unit charge on another is equal to the potential of the second due only to a unit charge on the first.* For example, a unit charge on a sphere produces a potential $1/\kappa r$ at an outside point distant r from the center. Hence a unit charge at this point raises the uncharged sphere to a potential $1/\kappa r$. Another proof of (19–3) will be found in article 21.

It is often convenient to express the charges in terms of the potentials. The solution of (19–1) for the charges gives

$$\left.\begin{aligned} Q_1 &= c_{11}V_1 + c_{12}V_2 + c_{13}V_3, \\ Q_2 &= c_{21}V_1 + c_{22}V_2 + c_{23}V_3, \\ Q_3 &= c_{31}V_1 + c_{32}V_2 + c_{33}V_3, \end{aligned}\right\} \tag{19–4}$$

where

$$\left.\begin{aligned} c_{11} &= \frac{p_{22}p_{33} - p_{23}^2}{\Delta}, & c_{12} = c_{21} &= \frac{p_{23}p_{31} - p_{12}p_{33}}{\Delta}, \\ c_{22} &= \frac{p_{33}p_{11} - p_{31}^2}{\Delta}, & c_{23} = c_{32} &= \frac{p_{31}p_{12} - p_{23}p_{11}}{\Delta}, \\ c_{33} &= \frac{p_{11}p_{22} - p_{12}^2}{\Delta}, & c_{31} = c_{13} &= \frac{p_{12}p_{23} - p_{31}p_{22}}{\Delta}, \end{aligned}\right\} \tag{19–5}$$

and

$$\Delta \equiv p_{11}p_{22}p_{33} + 2p_{12}p_{23}p_{31} - p_{11}p_{23}^2 - p_{22}p_{31}^2 - p_{33}p_{12}^2.$$

The quantities c_{11}, c_{22}, c_{33} are called *coefficients of capacity*, and the constants c_{12}, c_{23}, c_{31} *coefficients of induction*. If all conductors but one are kept at zero potential, a positive charge on that one produces a positive potential; so the coefficients of capacity are all positive. On the other hand, the charges induced on the other conductors are negative, or in special cases, zero; so the coefficients of induction are negative, or zero.

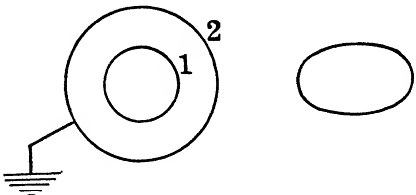


FIG. 40

The theory of *electric shielding* follows at once from (19–4). Let conductor (2) (Fig. 40) completely surround (1) and let $V_2 = 0$. Then

$$\begin{aligned} Q_1 &= c_{11}V_1 + c_{13}V_3, \\ Q_2 &= c_{21}V_1 + c_{23}V_3, \\ Q_3 &= c_{31}V_1 + c_{33}V_3. \end{aligned}$$

Since $V_1 = V_2 = 0$ if $Q_1 = 0$ (art. 11) regardless of V_3, it is evident that $c_{13} = 0$. In consequence

$$Q_1 = c_{11}V_1, \qquad Q_3 = c_{33}V_3, \tag{19–6}$$

so that the electric condition of (1) is entirely independent of (3) and *vice versa*. Thus a piece of electrostatic apparatus may be protected from external disturbances by enclosing it in a grounded metal box, a matter of great importance in making accurate measurements.

Problem 19a. Show that (*a*) $c_{11} + c_{12} + c_{13} \geq 0$, (*b*) all coefficients of potential are positive and $p_{11} - p_{12} \geq 0$.

Problem 19b. By solving (19–4) for the potentials find the p's in terms of the c's.

Problem 19c. Calculate the c's for the concentric spheres shown in Fig. 37. What change will be made by the presence of a third conductor as in Fig. 40?

$$\text{Ans. } c_{11} = \frac{\kappa ab}{b-a}, \quad c_{12} = -\frac{\kappa ab}{b-a}, \quad c_{22} = \frac{\kappa ab}{b-a} + b'.$$

20. Dielectric Absorption. — If dielectrics were perfect insulators, as we have tacitly assumed in the previous article, and if they had no internal resistance to polarization, a condenser would assume its full charge instantly on the application of a potential difference, and discharge instantly on being short-circuited. But experiment shows that with certain dielectrics a condenser continues to absorb charge for many minutes, and discharges in a similar manner. In fact, if such a condenser is short-circuited at regular intervals, being allowed to stand idle in between, repeated sparks of successively smaller intensity may be drawn from it. In measurement it is the first rush of charge or discharge, completed in a fraction of a second, that determines the *geometric capacity* of the condenser.

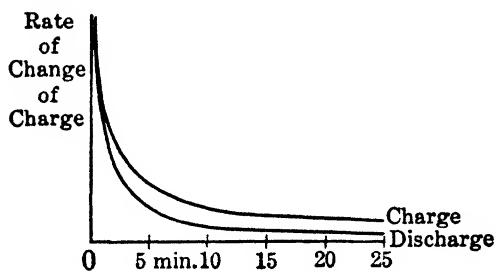


FIG. 41

The behavior of a typical condenser both while charging and while discharging is illustrated in Fig. 41, the ordinate indicating the rate of increase of charge in the case of the upper curve and the rate of decrease in the case of the lower. The two curves have the same form, but at any time they differ in value by a small amount. This is due to the dielectric's lack of perfection as an insulator; as long as the charging potential is applied there is a small but steady passage of electricity through the condenser. This is not real *absorption* of charge, of course, and on discharge it is not present. With liquid dielectrics or solids that have absorbed moisture the charge and the discharge curves sometimes differ in form as well as in magnitude.

In itself the slight conductivity of the dielectric is usually not important, but Maxwell showed that indirectly it may be responsible for the real absorption of charge, corresponding to the rapidly falling portion of the curves in Fig. 41. Maxwell's theory rests on the assumption that the dielectric layer is not homogeneous. As a very simple case let the insulating layer consist of two different dielectrics, mica and paper, say, of equal thickness (Fig. 42). Initially $D_1 = D_2$ by (14–5) so that $\kappa_1 E_1 = \kappa_2 E_2$, and the potential drops are, respectively,

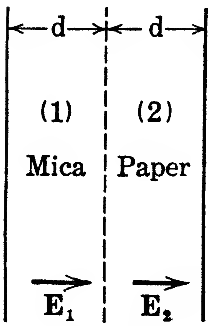


FIG. 42

$$E_1 d = \frac{\kappa_2}{\kappa_1 + \kappa_2} V_0, \qquad E_2 d = \frac{\kappa_1}{\kappa_1 + \kappa_2} V_0,$$

where V_0 is the total potential across the condenser. The charge passing through either substance is proportional to the potential drop across it. Thus $\dfrac{\kappa_2 G_1}{\kappa_1 + \kappa_2} V_0$ units of charge enter the dielectric layer per second, while $\dfrac{\kappa_1 G_2}{\kappa_1 + \kappa_2} V_0$ leave it, the G's representing the *conductivities* of the mica and the paper. In general

$$\frac{\kappa_1 G_2 - \kappa_2 G_1}{\kappa_1 + \kappa_2} V_0 \neq 0,$$

so that a charge collects on the boundary surface between (1) and (2). The presence of this charge decreases E_1 and increases E_2 until finally $G_1 E_1 d = G_2 E_2 d$ and there is no further accumulation of charge at the interface, but only a steady passage through the entire layer. When the condenser is short-circuited the process is reversed and we have, qualitatively at least, an explanation of Fig. 41. Of course, in practice we seldom have the ideal double-layer dielectric, but waxed paper approximates it, and it may be shown that any mixed or non-homogeneous dielectric should behave in the same way.

Maxwell's theory receives considerable support from experiment. Usually mixed dielectrics exhibit absorption and pure ones do not. On the other hand, the absorption curves may not be of the exact form predicted by the theory and absorption is occasionally found in seemingly pure homogeneous dielectrics. This suggests that there are other possible mechanisms of absorption, and a number have been proposed. One of the simplest is due to Décombe, who imagines that the atom or molecule offers a viscous resistance to polarization, so that when an electric field is applied the polarization rises to its final value slowly. This theory by itself seems incapable of explaining all the details of dielectric absorption, but it may well represent a contributing factor. In fact, it is very probable that the entire phenomenon of absorption is due to several mechanisms acting simultaneously.

21. Energy of Charged Systems. — To charge a system of conductors requires the expenditure of energy. This energy is stored electrostatically and becomes available again when the system is discharged. Let us build up a charge Q on an isolated conductor by bringing up infinitesimal charges dq through the surrounding medium. If q is the charge and v the potential at any time during this process, $q = Cv$ by (18–1). Similarly $Q = CV$, where Q and V represent the final charge and the final potential. Bringing up an infinitesimal charge from infinity does not change the potential by any finite amount; so the amount of work done, that is, the increase in the potential energy of the system during this process, is $v\,dq$. The total energy of the charged conductor is therefore given by

$$U = \int_0^Q v\,dq = \frac{1}{C}\int_0^Q q\,dq = \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2}QV = \frac{1}{2}CV^2. \quad (21\text{–}1)$$

Using e.s.u., this energy is in ergs, as usual.

With a condenser we may take the elements of charge from

the negative to the positive plate and

$$U = \int_0^Q (v_1 - v_2)dq = \frac{1}{C}\int_0^Q q\,dq = \frac{1}{2}\frac{Q^2}{C}$$
$$= \frac{1}{2}Q(V_1 - V_2) = \frac{1}{2}C(V_1 - V_2)^2. \quad (21\text{–}2)$$

By an extension of the above method we can calculate the energy of any number of charged conductors (Fig. 43). As before we build up the charge on each conductor in infinitesimal steps, but in such a way that every charge at any time is a given fraction of its final value. Thus, at any instant, $q_1 = \alpha\, Q_1$, $q_2 = \alpha\, Q_2$, $\cdots$ $q_k = \alpha\, Q_k$, where α lies between zero and unity. One step in the charging process must then consist of bringing up from infinity to (1) a charge $dq_1 = Q_1 d\alpha$, to (2) a charge $dq_2 = Q_2 d\alpha$ and so on. When all charges are multiplied by a given factor, the potentials are multiplied by the same factor; so, using the same notation as in (21–1), $v_1 = \alpha\, V_1$, $v_2 = \alpha\, V_2$, $\cdots$ $v_k = \alpha\, V_k$. The work done per step is $v_1 dq_1 + v_2 dq_2 + \cdots v_k dq_k$. Therefore

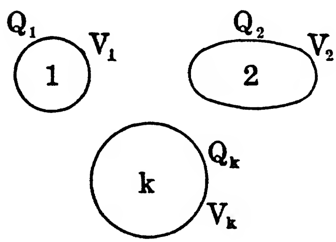


FIG. 43

$$U = \int_0^Q (v_1 dq_1 + v_2 dq_2 + \cdots v_k dq_k)$$
$$= (Q_1 V_1 + Q_2 V_2 + \cdots Q_k V_k)\int_0^1 \alpha\, d\alpha = \tfrac{1}{2}\sum_1^k Q_i V_i. \quad (21\text{–}3)$$

By means of the coefficients introduced in article 19 the energy may be expressed in terms of the charges or of the potentials. Restricting ourselves again for simplicity to three conductors,

$$\left.\begin{aligned} U_Q = \tfrac{1}{2}\{p_{11}Q_1^2 &+ p_{12}Q_1Q_2 + p_{13}Q_1Q_3 \\ + p_{21}Q_2Q_1 &+ p_{22}Q_2^2 + p_{23}Q_2Q_3 \\ + p_{31}Q_3Q_1 &+ p_{32}Q_3Q_2 + p_{33}Q_3^2\}, \end{aligned}\right\}(21\text{–}4)$$

$$\left.\begin{aligned} U_V = \tfrac{1}{2}\{c_{11}V_1^2 &+ c_{12}V_1V_2 + c_{13}V_1V_3 \\ + c_{21}V_2V_1 &+ c_{22}V_2^2 + c_{23}V_2V_3 \\ + c_{31}V_3V_1 &+ c_{32}V_3V_2 + c_{33}V_3^2\}. \end{aligned}\right\} \quad (21\text{–}5)$$

Now from (21–3)

$$2dU = (V_1dQ_1 + V_2dQ_2 + V_3dQ_3) + (Q_1dV_1 + Q_2dV_2 + Q_3dV_3).$$

But we know

$$dU = V_1dQ_1 + V_2dQ_2 + V_3dQ_3,$$

so that also

$$dU = Q_1dV_1 + Q_2dV_2 + Q_3dV_3.$$

Thus when the energy is expressed in terms of the Q's,

$$V_1 = \frac{\partial U_Q}{\partial Q_1}, \qquad V_2 = \frac{\partial U_Q}{\partial Q_2}, \qquad V_3 = \frac{\partial U_Q}{\partial Q_3}, \qquad (21\text{–}6)$$

and when in terms of the V's,

$$Q_1 = \frac{\partial U_V}{\partial V_1}, \qquad Q_2 = \frac{\partial U_V}{\partial V_2}, \qquad Q_3 = \frac{\partial U_V}{\partial V_3}. \qquad (21\text{–}7)$$

These differential relations are often useful. For example, from (21–6),

$$V_1 = p_{11}Q_1 + \tfrac{1}{2}(p_{12} + p_{21})Q_2 + \tfrac{1}{2}(p_{13} + p_{31})Q_3.$$

But

$$V_1 = p_{11}Q_1 + p_{12}Q_2 + p_{13}Q_3,$$

giving $p_{12} = p_{21}$, $p_{13} = p_{31}$; so that we have another proof of (19–3) that $p_{ij} = p_{ji}$. Similarly, using (21–7), we find $c_{ij} = c_{ji}$.

The amount of energy associated with a system of charges having been determined, the question of its location arises. This is evidently a somewhat arbitrary matter, since energy in itself is an intangible thing. When a dielectric is present, at least some of the energy is in the region between the charges, that is, in the field. It seems desirable therefore to distribute all the energy through the field. To see how this may be done, consider a single tube of electric displacement. This will begin on a free charge $1/4\pi$ and end on a free charge $-1/4\pi$. If the potentials at the ends of the tube are V_1 and V_2 respectively,

the contribution of the two charges to the energy of the entire system is, by (21–3),

$$\frac{1}{2}\left(\frac{1}{4\pi}\right)V_1 + \frac{1}{2}\left(-\frac{1}{4\pi}\right)V_2 = \frac{1}{8\pi}(V_1 - V_2) = \frac{1}{8\pi}\int_1^2 E\,dl, \quad (21\text{–}8)$$

where dl is an element of the path along the tube from (1) to (2). This energy *of* the tube we may distribute *along* the tube as we please. The simplest way is at the rate of $E/8\pi$ ergs per centimeter at every point along the tube, as this leads directly to (21–8) for the total energy of the tube. Since there are D tubes of displacement per square centimeter of cross-section (art. 14), energy is distributed through the field at the rate of $DE/8\pi$ ergs per cubic centimeter. Replacing D by κE for an isotropic dielectric,

$$U = \int_\tau \frac{\kappa E^2}{8\pi}\,d\tau, \quad (21\text{–}9)$$

the volume integral being taken through all space. The arbitrary nature of the distribution is quite evident from (21–9); for we may add to $\kappa E^2/8\pi$ any quantity whose integral through all space is zero, and still obtain the correct result. This distribution is, however, more plausible than others. For in isotropic media

$$D = E + 4\pi P,$$

so that

$$\frac{DE}{8\pi} = \frac{1}{8\pi}E^2 + \frac{1}{2}EP, \quad (21\text{–}10)$$

the last term being the energy per cubic centimeter assignable to the dielectric proper.

Now if an electric doublet of moment $p = ql$ is created by a field E, the work performed is $qEl = Ep$. Therefore the work done in increasing the electric moment per unit volume from P to $P + dP$ is given by

$$EdP = \epsilon EdE,$$

and the total energy stored in the dielectric per unit volume by

$$\int_0^E \epsilon EdE = \tfrac{1}{2}\epsilon E^2 = \tfrac{1}{2}EP,$$

which is exactly the amount specified by the arbitrary distribution above.

The concept of energy in the field is an illuminating one, particularly in the study of forces on conductors and dielectrics. Use will be made of it in the following article.

Problem 21a. A sphere whose radius is 20 cm is charged to a potential of 100 units; find its energy. It is then connected by a long thin wire with an equal uncharged sphere; find the energy of the system. What has become of the rest of the energy? Ans. 100,000 erg, 50,000 erg.

Problem 21b. Two condensers of capacities C_1 and C_2 have charges Q_1 and Q_2 respectively. Calculate the amount of energy dissipated when they are connected in parallel. Ans. $\dfrac{(Q_1C_2 - Q_2C_1)^2}{2C_1C_2(C_1 + C_2)}$.

Problem 21c. A condenser ($\kappa = 1$) composed of concentric spheres (Fig. 37) is so constructed that the outer sphere can be separated and the inner sphere removed without changing the charges on either. The radius of the inner sphere is a, of the outer b, and the charges are Q and $-Q$ respectively. Let the inner sphere be removed, the outer restored to its original form, and the two separated by a great distance. Calculate the increase of energy during this process. What is the source of it? Ans. $\dfrac{Q^2}{b}$.

22. Forces and Torques. — When a system changes its configuration under its own forces work is done, which can only be at the expense of its potential energy, if there is no external supply of energy. Thus, let conductor (1) move a distance $d\xi$ under a force F, the charges on all conductors in the system being kept constant. Then

$$dU_Q = -Fd\xi,$$

$$F = -\frac{\partial U_Q}{\partial \xi}. \tag{22–1}$$

Here we are differentiating the p's in the expression (21–4), the Q's remaining constant.

It often happens that the potentials are constant instead of the charges. For a given configuration of the system the forces are the same, of course, whether charges or potentials are kept

constant. However, the work done in a change of configuration no longer comes from the potential energy; so F must be calculated differently. For simplicity we limit ourselves to one movable and one fixed conductor, but the number of fixed conductors does not affect the result. We may imagine that the displacement $d\xi$ with constant V's is performed in two steps, a motion with constant Q's, followed by an addition of charges sufficient to restore the potentials to their initial values. If the energy added in the latter step is δU,

$$dU_V = dU_Q + \delta U, \tag{22-2}$$

where dU_V is the desired energy increase due to displacement with the V's constant. During the first step the potentials change by

$$dV_1 = \left(\frac{\partial p_{11}}{\partial \xi} Q_1 + \frac{\partial p_{12}}{\partial \xi} Q_2\right) d\xi,$$

$$dV_2 = \left(\frac{\partial p_{21}}{\partial \xi} Q_1 + \frac{\partial p_{22}}{\partial \xi} Q_2\right) d\xi.$$

Hence in the second step we must add charges dQ_1 and dQ_2 given by

$$p_{11}dQ_1 + p_{12}dQ_2 = -\left(\frac{\partial p_{11}}{\partial \xi} Q_1 + \frac{\partial p_{12}}{\partial \xi} Q_2\right) d\xi,$$

$$p_{21}dQ_1 + p_{22}dQ_2 = -\left(\frac{\partial p_{21}}{\partial \xi} Q_1 + \frac{\partial p_{22}}{\partial \xi} Q_2\right) d\xi,$$

which will cause the energy to increase by

$$\begin{aligned}\delta U &= V_1 dQ_1 + V_2 dQ_2 \\ &= (p_{11}dQ_1 + p_{21}dQ_2)Q_1 + (p_{12}dQ_1 + p_{22}dQ_2)Q_2 \\ &= -\left(\frac{\partial p_{11}}{\partial \xi} Q_1^2 + 2\frac{\partial p_{12}}{\partial \xi} Q_1 Q_2 + \frac{\partial p_{22}}{\partial \xi} Q_2^2\right) d\xi,\end{aligned}$$

since $p_{12} = p_{21}$. But

$$U_Q = \tfrac{1}{2}(p_{11}Q_1^2 + 2p_{12}Q_1Q_2 + p_{22}Q_2^2),$$

and therefore

$$\delta U = -2dU_Q.$$

This gives

$$dU_V = -\,dU_Q = Fd\xi,$$

and

$$F = \frac{\partial U_V}{\partial \xi}\,. \tag{22–3}$$

Here we are differentiating the c's in the expression (21–5), the V's being held constant. This is an interesting result; for it shows that in motions with constant potentials the potential energy *increases* by an amount exactly equal to the mechanical work done. In other words, the sources of the constant potentials supply double energy to the system.

Torques are calculated by differentiating with respect to an angle. For constant charges the torque is given by

$$L = -\frac{\partial U_Q}{\partial \theta}, \tag{22–4}$$

and for constant potentials by

$$L = \frac{\partial U_V}{\partial \theta}\,. \tag{22–5}$$

A piece of dielectric in an electric field experiences forces, in general, just as a conductor does, and these are calculated according to the same principles. It is usually more convenient, however, to use the distributed energy function (21–9), rather than (21–3). For example, the energy per unit volume may be written $D^2/8\pi\kappa$, and as the tubes of displacement are continuous across the dielectric boundary the potential energy of the system will be least when the greatest possible number of tubes are included in the dielectric. Therefore, the forces on a dielectric are such as to move it from a weaker to a stronger part of the field.

— *Problem 22a.* A parallel plate condenser of length l, width b and plate separation d has the space between the plates filled by a slab of dielectric whose constant is κ. This slab is withdrawn in the direction of its length until only a length x remains between the plates. The potential difference V of the plates is maintained constant. Neglecting edge effects calculate the force tending to restore the slab to its original position. Ans. $\dfrac{(\kappa - 1)V^2}{8\pi d}\,b$.

23. Electrostatic Instruments. — Mention has already been made of the electroscope. This is a sensitive instrument but not one adapted to accurate quantitative measurements, except in special cases. There are, however, accurate potential measuring instruments, known as *electrometers*.

The Absolute Electrometer. — This instrument (Fig. 44) is essentially a parallel plate condenser with a guard ring, the guarded plate being suspended from a balance arm, so that the force F on it is measurable in terms of weight. The distance between the plates and the mass in the pan are adjusted until

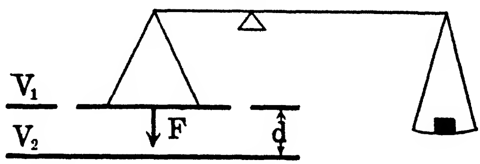


FIG. 44

the movable plate hangs exactly in the plane of the guard ring. Then, if A' is the area of the plate plus one-half the gap, as in (18–10),

$$F = 2\pi\sigma^2 A'.$$

But

$$\frac{V_1 - V_2}{d} = E = 4\pi\sigma,$$

so that

$$F = 2\pi\left(\frac{V_1 - V_2}{4\pi d}\right)^2 A'$$

and

$$V_1 - V_2 = d\sqrt{\frac{8\pi F}{A'}}. \qquad (23\text{–}1)$$

As F is known directly in terms of fundamental quantities, the instrument does not require calibration; hence the designation "absolute."

The Quadrant Electrometer. — This is the most useful of the electrostatic instruments. It combines high sensitivity with great accuracy, and is adaptable to a great variety of uses. It

consists of a small metal pill-box cut into quadrants, shown both in plan and in section in Fig. 45. Within is a metal vane, known as the *needle*, which is suspended by a fine torsion fibre. Opposite quadrants are connected by a wire, the potentials of the two pairs being respectively V_1 and V_2. The needle is maintained at some known potential V_3. If it assumes a position symmetrical with the quadrants when both pairs are at the same potential, it will turn through some angle θ when different potentials are applied to the two. When the constants of the instrument are known, $V_1 - V_2$ may be determined from θ. In practice this angle is determined by means of a small mirror mounted on the needle suspension, and an external telescope and scale. The instrument has a removable metal case which serves as an electrostatic shield. The details of construction are shown in Fig. 46, the case having been removed.

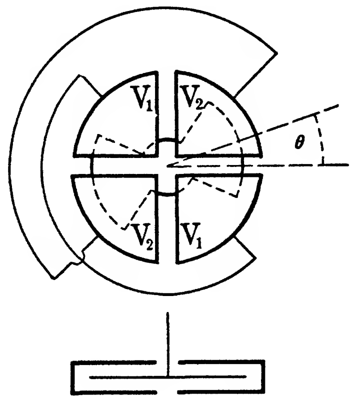


FIG. 45

The complete theory of the quadrant electrometer is somewhat elaborate, but a sufficiently accurate result is obtained by an approximate treatment. Since the quadrants are edge to edge, the direct capacity between them is small compared to the capacity between quadrants and needle. Hence the energy of the system may be expressed approximately as

$$U_V = \tfrac{1}{2}C_1(V_3 - V_1)^2 + \tfrac{1}{2}C_2(V_3 - V_2)^2 + U_0, \quad (23\text{-}2)$$

where C_1 is the capacity of quadrants (1) relative to the needle, neglecting the existence of quadrants (2). C_2 refers similarly to quadrants (2) and the needle, and U_0 is a term independent of the needle, involving the capacity of the quadrants to ground. C_1 and C_2 are evidently functions of θ. The torque on the needle

is, by (22–5),

$$L = \frac{\partial U_V}{\partial \theta} = \frac{1}{2}\frac{\partial C_1}{\partial \theta}(V_3 - V_1)^2 + \frac{1}{2}\frac{\partial C_2}{\partial \theta}(V_3 - V_2)^2.$$

This torque is balanced by that of the fibre, which, as it is proportional to θ, may be written $a\theta$. Also, since the needle

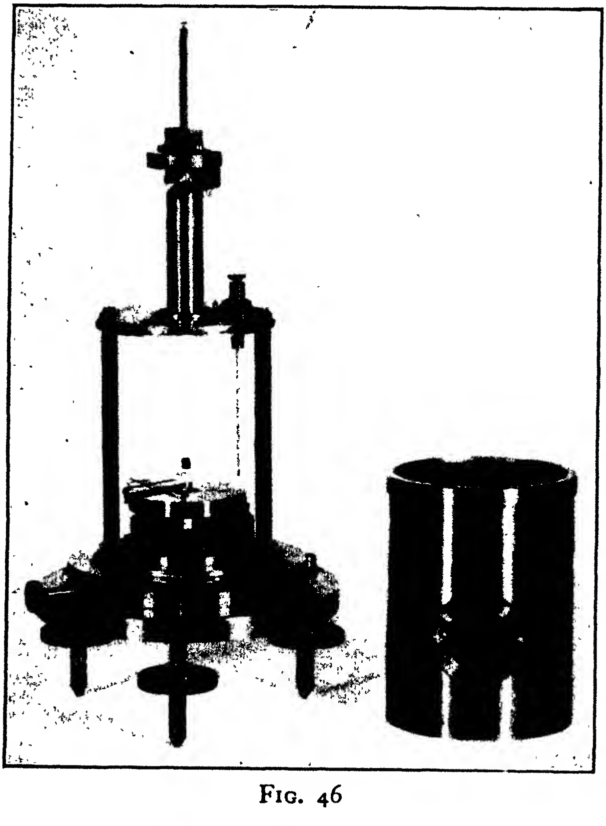


Fig. 46

turns out of one pair of quadrants as much as it turns into the other,

$$-\frac{\partial C_1}{\partial \theta} = \frac{\partial C_2}{\partial \theta} = b,$$

another constant. Therefore

$$a\theta = -\tfrac{1}{2}b(V_3 - V_1)^2 + \tfrac{1}{2}b(V_3 - V_2)^2,$$

or

$$\theta = k(V_1 - V_2)\left\{V_3 - \frac{V_1 + V_2}{2}\right\}, \qquad (23\text{-}3)$$

where k is a constant to be determined by calibration.

There are two methods of use. In the *heterostatic* method V_3 is large compared to V_1 and V_2 and is held constant. Then (23-3) reduces to

$$\theta = kV_3(V_1 - V_2), \qquad (23\text{-}4)$$

so that the deflection is proportional to $(V_1 - V_2)$.

In the *idiostatic* method the needle is connected to one pair of quadrants, say (1). This gives

$$\theta = \tfrac{1}{2}k(V_1 - V_2)^2. \qquad (23\text{-}5)$$

This method is less sensitive than the first, but permits the measurement of alternating potentials, since the deflection is always in one direction.

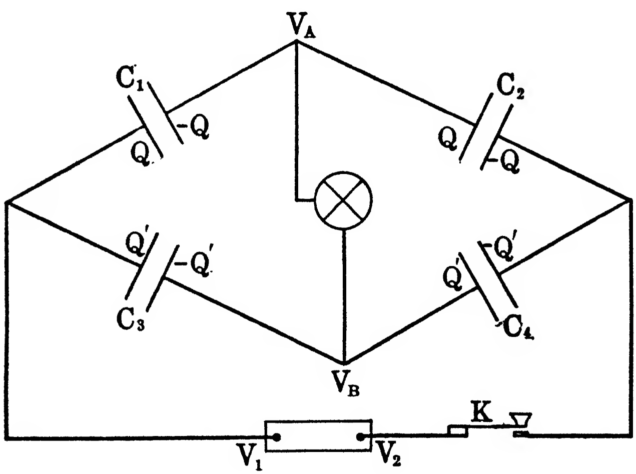


FIG. 47

While the quadrant electrometer is not an absolute instrument, it is much more sensitive than the absolute electrometer. Used heterostatically it can measure $(10)^{-6}$ e.s.u. of potential without difficulty. With special precautions an even greater

sensitivity is obtainable. In precise measurements it is usually used as a *null* instrument. An example of such use is the *capacity bridge*, illustrated in Fig. 47. C_1 is an unknown capacity; C_2, C_3, C_4 are known. C_2, being variable, is adjusted until the electrometer shows no deflection when a potential difference $V_1 - V_2$ is placed across the bridge by closing the key K. Under this condition $V_A = V_B$ and the quadrants are uncharged, so that C_1 has the same charge as C_2, and C_3 the same as C_4. Then

$$V_1 - V_A = \frac{Q}{C_1}, \qquad V_A - V_2 = \frac{Q}{C_2},$$

$$V_1 - V_B = \frac{Q'}{C_3}, \qquad V_B - V_2 = \frac{Q'}{C_4}.$$

Therefore

$$C_1(V_1 - V_A) = C_2(V_A - V_2),$$
$$C_3(V_1 - V_B) = C_4(V_B - V_2),$$

and as $V_A = V_B$,

$$\frac{C_1}{C_2} = \frac{C_3}{C_4}. \qquad (23\text{-}4)$$

24. Determination of the Dielectric Constant. — Dielectric constants were first measured by Faraday, by means of two geometrically identical spherical condensers, one of which had air as a dielectric, the other some solid such as sulphur. He found the ratio of the capacities, which is κ, by charging one condenser to a known potential, allowing it to share its charge with the other, and then determining the final potential. Boltzmann later employed a similar scheme using a parallel plate condenser into which a slab of dielectric could be introduced. These early measurements were all subject to error on account of dielectric absorption (art. 20). This error is now avoided by means of quick acting switches and other special devices.

A table is appended giving values of κ for the more common dielectrics. The gases are at 0° C and under a pressure of one

atmosphere. It is worthy of note that for practically all solids κ lies between one and ten.

Substance	κ
Solids:	
Ebonite	2.7–2.9
Glass	5.0–10.0
Mica	5.7–6.5
Paper	2.0–2.5
Paraffin	2.3
Quartz	4.5
Shellac	3.1–3.7
Sulphur	3.6–4.2
Liquids:	
Castor oil	4.7
Kerosene	4.6–4.8
Turpentine	2.2
Water	81.0
Gases:	
Air	1.000588
Hydrogen	1.000264
Carbon Dioxide	1.000985

25. Practical Units. — Because of the inconvenient magnitude of the fundamental electrostatic units, a secondary or *practical* set of units is commonly used. The practical unit of

Charge	is the	*coulomb.*	It is $3(10)^9$ e.s.u.
Potential	" "	*volt.*	" " 1/300 e.s.u.
Capacity	" "	*farad.*	" " $9\ (10)^{11}$ e.s.u.
		microfarad.	" " $9\ (10)^5$ e.s.u.
Electric Intensity	" "	*volt per cm.*	" " 1/300 e.s.u.
Energy	" "	*joule.*	" " $(10)^7$ ergs.

The dielectric constant is the same in either system of units.

Observe that while some formulas have the same form in practical units as in e.s.u., others involve numerical constants on transformation. Thus, using the subscript p to indicate

practical units,

$$Q_p = C_p V_p, \qquad E_p l = V_p,$$

but

$$F = 9\,(10)^{18}\,\frac{q_p q_p'}{\kappa r^2}, \qquad V_p = 9\,(10)^{11}\,\frac{q_p}{\kappa r}.$$

A more detailed discussion of units, and transformation from one system to another will be found in article 50 and in Chapter XII.

CHAPTER III

SOLUTION OF ELECTROSTATIC PROBLEMS

26. Poisson's and Laplace's Equations. — We shall develop now the differential equations which must be satisfied by the potential in an isotropic medium in which are immersed free charges of density ρ per unit volume. In most problems of importance the free charges are located on the surfaces of conductors surrounded by a dielectric and the solution of the problem consists in finding the potential at all points in the dielectric in terms of the assigned potentials of the conducting surfaces. Once the potential is known as a function of the coordinates the electric intensity is obtained at once from (7–7) or (7–8) and the charge per unit area on the surfaces of the conductors is given in terms of the electric intensity immediately outside by (15–10).

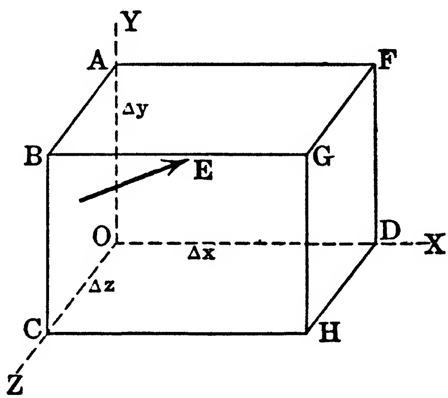


FIG. 48

Consider an infinitesimal rectangular parallelopiped (Fig. 48) with its edges parallel to the coordinate axes. The flux of displacement inward through the face $OABC$ is

$$\kappa E_x \Delta y \Delta z,$$

and that outward through the face $DFGH$ is

$$\kappa \left(E_x + \frac{\partial E_x}{\partial x} \Delta x \right) \Delta y \Delta z.$$

The net flux outward through these two faces perpendicular to the X axis is the difference of the two expressions, that is,

$$\kappa \frac{\partial E_x}{\partial x} \Delta x \Delta y \Delta z.$$

Adding similar expressions for the outward flux through the pairs of faces perpendicular to the Y and Z axes we have for the total outward flux

$$\kappa \left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \right) \Delta x \Delta y \Delta z.$$

Gauss' law (14–4) requires this to be equal to 4π times the free charge contained in the parallelopiped. So if the density of free charge is denoted by ρ,

$$\kappa \left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \right) \Delta x \Delta y \Delta z = 4\pi\rho \Delta x \Delta y \Delta z.$$

The components of electric intensity are given in terms of the potential by (7–7). Using these relations and dividing through by $\kappa \Delta x \Delta y \Delta z$ the equation above becomes

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -\frac{4\pi\rho}{\kappa}. \qquad (26\text{–}1)$$

This equation is known as *Poisson's equation.* In a region where no free charges are present $\rho = 0$ and (26–1) reduces to *Laplace's equation*,

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0. \qquad (26\text{–}2)$$

If, now, we are given a set of conductors at potentials V_1, V_2, $\cdots V_k$ the determination of the potential in the region between and surrounding the conductors resolves itself into the problem of finding a solution of (26–2) which reduces to V_1 at the surface of the first conductor, V_2 at the surface of the second and so on. Such a solution may be shown to be *unique* and therefore to represent the *only* solution of the problem under consideration. A solution which represents the field outside the conductors,

however, does not apply to their interiors, for there is charge on the surface of each conductor and therefore Laplace's equation does not hold across the surface. In fact the potential inside a conductor is always constant, as was shown in article 11.

In (26–2) we have Laplace's equation expressed in rectangular coordinates. In many problems other coordinates are more suitable. We shall therefore deduce this equation again in general orthogonal curvilinear coordinates ξ, η, ζ. The coordinate surfaces consist of the three families $\xi(x, y, z) =$ constant, $\eta(x, y, z) =$ constant, $\zeta(x, y, z) =$ constant, intersecting one another at right angles. Consider a small volume (Fig. 49) bounded by coordinate surfaces. Let E_ξ, E_η, E_ζ be the components of the electric intensity in the directions of increasing ξ, η, ζ respectively, and λ, μ, ν the functions of ξ, η, ζ by which $\Delta\xi$, $\Delta\eta$, $\Delta\zeta$ respectively must be multiplied in order to obtain the distances corresponding to these increments in the coordinates. Then the area of $OABC$ is $\mu\nu\Delta\eta\Delta\zeta$ and the flux through it is

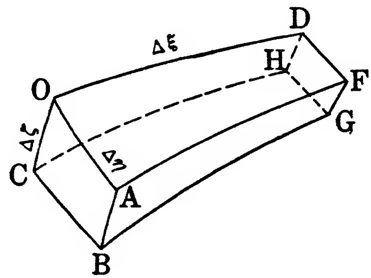


FIG. 49

$$\kappa E_\xi \mu\nu\Delta\eta\Delta\zeta.$$

Similarly the flux through $DFGH$ is

$$\kappa\left\{E_\xi\mu\nu + \frac{\partial}{\partial\xi}(E_\xi\mu\nu)\Delta\xi\right\}\Delta\eta\Delta\zeta.$$

Subtracting, the net outward flux through the two surfaces is seen to be

$$\kappa\frac{\partial}{\partial\xi}(E_\xi\mu\nu)\Delta\xi\Delta\eta\Delta\zeta.$$

Adding similar terms for the flux through the remaining two pairs of surfaces, equating the sum to zero and dividing by the

common factor $\kappa\Delta\xi\Delta\eta\Delta\zeta$, we get

$$\frac{\partial}{\partial\xi}(\mu\nu E_\xi) + \frac{\partial}{\partial\eta}(\nu\lambda E_\eta) + \frac{\partial}{\partial\zeta}(\lambda\mu E_\zeta) = 0.$$

In accord with (7–6),

$$E_\xi = -\frac{\partial V}{\lambda\partial\xi}, \qquad E_\eta = -\frac{\partial V}{\mu\partial\eta}, \qquad E_\zeta = -\frac{\partial V}{\nu\partial\zeta}.$$

So Laplace's equation takes the form

$$\frac{\partial}{\partial\xi}\left(\frac{\mu\nu}{\lambda}\frac{\partial V}{\partial\xi}\right) + \frac{\partial}{\partial\eta}\left(\frac{\nu\lambda}{\mu}\frac{\partial V}{\partial\eta}\right) + \frac{\partial}{\partial\zeta}\left(\frac{\lambda\mu}{\nu}\frac{\partial V}{\partial\zeta}\right) = 0. \qquad (26\text{–}3)$$

In the case of spherical coordinates ξ, η, ζ become r, θ, ϕ and λ, μ, ν are 1, r, $r\sin\theta$ respectively. Hence Laplace's equation in spherical coordinates is

$$\frac{\partial}{\partial r}\left(r^2\sin\theta\frac{\partial V}{\partial r}\right) + \frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial V}{\partial\theta}\right) + \frac{\partial}{\partial\phi}\left(\frac{1}{\sin\theta}\frac{\partial V}{\partial\phi}\right) = 0.$$

Dividing by $\sin\theta$ this may be put in the better form,

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial V}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 V}{\partial\phi^2} = 0. \qquad (26\text{–}4)$$

In cylindrical coordinates ξ, η, ζ become r, θ, z and λ, μ, ν assume the values 1, r, 1. Therefore Laplace's equation becomes

$$\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right) + \frac{\partial}{\partial\theta}\left(\frac{1}{r}\frac{\partial V}{\partial\theta}\right) + \frac{\partial}{\partial z}\left(r\frac{\partial V}{\partial z}\right) = 0,$$

or

$$r\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right) + \frac{\partial^2 V}{\partial\theta^2} + r^2\frac{\partial^2 V}{\partial z^2} = 0. \qquad (26\text{–}5)$$

To obtain the charge density σ per unit area on the surface of a conductor when we know the potential function V outside the conductor we have, from (15–10) and (7–6),

$$\sigma = \frac{\kappa}{4\pi}E = -\frac{\kappa}{4\pi}\frac{\partial V}{\partial n}, \qquad (26\text{–}6)$$

where $\dfrac{\partial V}{\partial n}$ is the space derivative of V along the outward normal n to the surface of the conductor.

Problem 26a. Obtain (26–4) directly from (26–2) by transforming the coordinates according to the scheme $x = r \sin\theta \cos\phi$, $y = r \sin\theta \sin\phi$, $z = r\cos\theta$.

Problem 26b. Obtain (26–5) directly from (26–2) by means of the transformations $x = r\cos\theta$, $y = r\sin\theta$.

27. Solutions of Laplace's Equation. — In solving Laplace's equation the two theorems that follow are of importance.

Theorem 1. If $V_1, V_2, \cdots V_k$ are solutions of Laplace equation, then

$$V = A_1V_1 + A_2V_2 + \cdots A_kV_k$$

is also a solution, where the A's are arbitrary constants. This theorem is proved at once by substituting the expression above in (26–2).

Theorem 2. If V is a solution of Laplace's equation, then

$$\frac{\partial V}{\partial x}, \quad \frac{\partial V}{\partial y}, \quad \frac{\partial V}{\partial z}, \quad \frac{\partial^2 V}{\partial x^2}, \quad \frac{\partial^2 V}{\partial x \partial y}, \quad \text{etc.},$$

and in fact all partial derivatives of V with respect to one or more of the rectangular coordinates x, y, z (but not with respect to spherical coordinates) are solutions. For if we differentiate (26–2) partially with respect to x we have

$$\frac{\partial^2}{\partial x^2}\left(\frac{\partial V}{\partial x}\right) + \frac{\partial^2}{\partial y^2}\left(\frac{\partial V}{\partial x}\right) + \frac{\partial^2}{\partial z^2}\left(\frac{\partial V}{\partial x}\right) = 0.$$

Spherical Coordinates.—In many electrostatic problems we are concerned with conducting spheres and spherical coordinates are indicated for the solution. We shall limit our discussion to those cases in which the potential is a function of the radius vector r and the polar angle θ alone. Then the last term in (26–4) drops out and if we put μ for $\cos\theta$ Laplace's equation becomes

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) + \frac{\partial}{\partial \mu}\left\{(1-\mu^2)\frac{\partial V}{\partial \mu}\right\} = 0. \qquad (27\text{–}1)$$

This differential equation has a solution of the form

$$V = r^n P_n,$$

where P_n is a function of μ alone. Substituting in (27–1) we find that P_n must satisfy the differential equation

$$\frac{d}{d\mu}\left\{(1-\mu^2)\frac{dP_n}{d\mu}\right\} + n(n+1)P_n = 0, \qquad (27\text{–}2)$$

which is known as *Legendre's equation*. Now if we replace n in this equation by $-(n+1)$ the coefficient of the last term becomes

$$\{-(n+1)\}\{-(n+1)+1\} = n(n+1).$$

Therefore Legendre's equation for $P_{-(n+1)}$ is the same as that for P_n. Consequently P_n and $P_{-(n+1)}$ are identical. Thus every P_n satisfying Legendre's equation provides us with *two* solutions of Laplace's equation, namely,

$$r^n P_n \quad \text{and} \quad \frac{P_n}{r^{n+1}}.$$

Now

$$V_0 = \frac{1}{r}, \qquad P_0 = 1, \qquad (27\text{–}3)$$

satisfies Laplace's equation, as is seen at once by substitution in (27–1). Let us take the X axis in the direction of the polar axis of the spherical coordinates, so that $x = r\cos\theta$. We can get a second solution from (27–3) by differentiating partially with respect to x. Remembering that $\dfrac{\partial r}{\partial x} = \dfrac{x}{r} = \cos\theta$,

$$V_1 = \frac{x}{r^3} = \frac{\cos\theta}{r^2}, \qquad P_1 = \cos\theta, \qquad (27\text{–}4)$$

where we have dropped a constant factor since any solution may be multiplied by an arbitrary constant.

Differentiating again with respect to x to get a third solution,

$$V_2 = \frac{3x^2}{r^5} - \frac{1}{r^3} = \frac{3\cos^2\theta - 1}{r^3}, \qquad P_2 = 3\cos^2\theta - 1, \qquad (27\text{–}5)$$

and repeating the process for a fourth solution,

$$V_3 = \frac{5x^3}{r^7} - \frac{3x}{r^5} = \frac{5\cos^3\theta - 3\cos\theta}{r^4},$$

$$P_3 = 5\cos^3\theta - 3\cos\theta. \qquad (27\text{–}6)$$

The solutions we have obtained are of the form P_n/r^{n+1} but as we have found the P_n's we can write down at once solutions of the form $r^n P_n$. Table I contains the solutions which we have deduced and which suffice for the applications which we shall have occasion to make. The functions listed are known as *zonal harmonics.*

Table I

Zonal Harmonics

$1,$	$\frac{1}{r},$
$r \cos \theta,$	$\frac{1}{r^2} \cos \theta,$
$r^2 (3 \cos^2 \theta - 1),$	$\frac{1}{r^3} (3 \cos^2 \theta - 1),$
$r^3 (5 \cos^3 \theta - 3 \cos \theta).$	$\frac{1}{r^4} (5 \cos^3 \theta - 3 \cos \theta).$

Cylindrical Coordinates.—In some problems—such as those having to do with a long straight wire—the potential is not a function of one of the rectangular coordinates, say z, and it is convenient to use polar coordinates in the plane perpendicular to the Z axis. In this case the last term in (26–5) disappears and Laplace's equation reduces to

$$r \frac{\partial}{\partial r}\left(r \frac{\partial V}{\partial r} \right) + \frac{\partial^2 V}{\partial \theta^2} = 0. \qquad (27\text{–}7)$$

Substituting

$$V = r^n C_n,$$

where C_n is a function of θ only, we see that C_n must satisfy the equation

$$\frac{d^2 C_n}{d\theta^2} + n^2 C_n = 0. \qquad (27\text{–}8)$$

Evidently the differential equation for C_{-n} is the same as that for C_n. Therefore C_{-n} and C_n are identical, and every C_n satisfying (27–8) leads to two solutions of Laplace's equation,

namely,

$$r^n C_n \quad \text{and} \quad \frac{C_n}{r^n}.$$

The simplest solution of (27–7) other than a constant is

$$V_0 = \log r, \tag{27–9}$$

the logarithm, unless otherwise noted, always being taken to base e. To get other solutions we will differentiate partially with respect to $x = r \cos \theta$. Then we find

$$V_1 = \frac{x}{r^2} = \frac{\cos \theta}{r}, \qquad C_1 = \cos \theta; \tag{27–10}$$

$$V_2 = \frac{2x^2}{r^4} - \frac{1}{r^2} = \frac{\cos 2\theta}{r^2}, \qquad C_2 = \cos 2\theta; \tag{27–11}$$

$$V_3 = \frac{4x^3}{r^6} - \frac{3x}{r^4} = \frac{\cos 3\theta}{r^3}, \qquad C_3 = \cos 3\theta. \tag{27–12}$$

Although the solutions we have obtained are of the form C_n/r^n we can write down at once solutions of the form $r^n C_n$ since we know the C_n's. It may be noted that if we had differentiated (27–9) with respect to $y = r \sin \theta$ instead of $x = r \cos \theta$ the cosines would have been replaced by sines. In fact the complete solution of (27–8) is readily seen to be

$$C_n = A_n \sin n\theta + B_n \cos n\theta.$$

Table II contains the *cylindrical harmonics* which we have found.

Table II

Cylindrical Harmonics

1;	$\log r$;
$r \cos \theta,\ r \sin \theta$;	$\dfrac{\cos \theta}{r},\ \dfrac{\sin \theta}{r}$;
$r^2 \cos 2\theta,\ r^2 \sin 2\theta$;	$\dfrac{\cos 2\theta}{r^2},\ \dfrac{\sin 2\theta}{r^2}$;
$r^3 \cos 3\theta,\ r^3 \sin 3\theta$.	$\dfrac{\cos 3\theta}{r^3},\ \dfrac{\sin 3\theta}{r^3}$.

Problem 27a. If R is the distance of a point P from the origin O, a the distance of a point Q from O, and r the distance of P from Q,

$$R = \sqrt{r^2 + 2ar \cos \theta + a^2},$$

where θ is the angle between OQ and QP. Expand the reciprocal of R by the binomial theorem both for the case $r < a$ and for the case $r > a$. Note that the successive terms in the expansion are zonal harmonics.

28. Sphere in Uniform Field. — We will use zonal harmonics to solve the problem of a sphere placed in a uniform electric field, treating first the case of a conducting sphere and second that of a dielectric sphere. In each case the problem consists in finding the potential V as a function of r and θ from which the components of electric intensity can be obtained by means of (7–8) and the charge density on the surface of the conducting sphere from (26–6).

Conducting Sphere in Uniform Field.—Consider an uncharged conducting sphere of radius a placed in a uniform field E_0. As potential is a relative quantity we may take the potential of the sphere as zero. While the sphere distorts the field in its neighborhood, as shown in Fig. 50, the field at a great distance

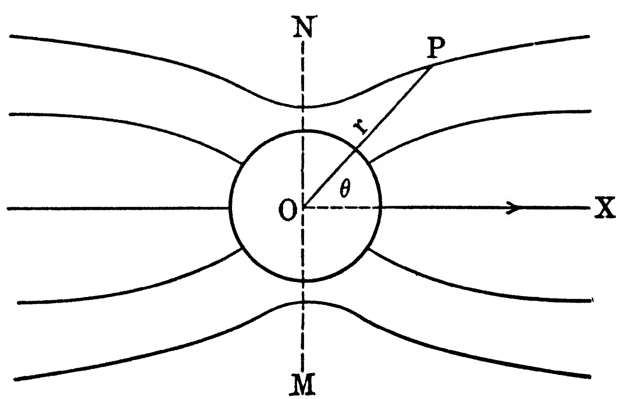


FIG. 50

retains its original uniform character. Therefore, if the origin is taken at the center of the sphere and the X axis in the direction

of E_0, the potential at a great distance from the sphere must be $-E_0x = -E_0r\cos\theta$. So we must look for a solution of Laplace's equation which satisfies the two boundary conditions

$$V = -E_0r\cos\theta \quad \text{for} \quad r = \infty,$$
$$V = 0 \qquad\qquad \text{for} \quad r = a.$$

To satisfy the first of these conditions the zonal harmonic $r\cos\theta$ is required and to satisfy the second another harmonic involving only the first power of $\cos\theta$ must be added. Therefore a glance at Table I shows that the potential at a point P must be of the form

$$V = Ar\cos\theta + \frac{B\cos\theta}{r^2},$$

where the arbitrary constants A and B are to be determined so as to satisfy the boundary conditions. When r is infinite the second term in V disappears, and the first boundary condition is satisfied if $A = -E_0$. To satisfy the second we put r equal to a and equate the coefficient of $\cos\theta$ to zero, getting $B = E_0a^3$. Hence the potential outside the sphere is

$$V = -\left(1 - \frac{a^3}{r^3}\right)E_0r\cos\theta. \tag{28-1}$$

Inside the sphere the potential is everywhere zero.

The components of electric intensity outside the sphere are

$$E_r = -\frac{\partial V}{\partial r} = \left(1 + 2\frac{a^3}{r^3}\right)E_0\cos\theta \tag{28-2}$$

along the radius vector, and

$$E_\theta = -\frac{\partial V}{r\partial\theta} = -\left(1 - \frac{a^3}{r^3}\right)E_0\sin\theta \tag{28-3}$$

at right angles to the radius vector in the direction of increasing θ.

As the radius vector is normal to the surface of the sphere the charge per unit area on its surface is

$$\sigma = -\frac{\kappa}{4\pi}\left(\frac{\partial V}{\partial r}\right)_{r=a} = \frac{3\kappa E_0\cos\theta}{4\pi}, \tag{28-4}$$

being positive on the right-hand hemisphere and negative on the left.

So far we have considered an uncharged sphere at zero potential. If the sphere has a charge Q we must add $Q/\kappa r$ to (28–1), getting

$$V = -\left(1 - \frac{a^3}{r^3}\right) E_0 r \cos\theta + \frac{Q}{\kappa r}. \tag{28–5}$$

The radial component of the electric intensity is increased by $Q/\kappa r^2$ and the charge per unit area by $Q/4\pi a^2$.

To return to the case of the uncharged sphere, it is clear from symmetry that the plane MN (Fig. 50) through the center of the sphere at right angles to X is an equipotential surface at the same potential as the sphere. Therefore we can replace the portion of this plane outside the sphere by a conducting surface, wiping out the field to the left and leaving that to the right unaltered. We have, then, the case of an infinite conducting plane with a hemispherical boss as illustrated in Fig. 51. The lines of force originate on positive charges on the surface of the plane instead of starting from positive charges at an infinite distance to the left. The potential is still given by (28–1) and the charge per unit area on the hemisphere by (28–4).

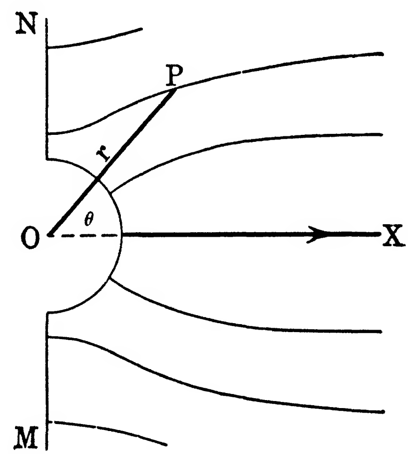


FIG. 51

To find the charge per unit area on the plane portion of the surface write (28–1) in the form

$$V = -\left(1 - \frac{a^3}{r^3}\right) E_0 x.$$

As the X axis is normal to the plane,

$$\sigma = -\frac{\kappa}{4\pi}\left(\frac{\partial V}{\partial x}\right)_{x=0} = \left(1 - \frac{a^3}{r^3}\right)\frac{\kappa E_0}{4\pi}. \qquad (28\text{–}6)$$

Inspection of (28–4) and (28–6) shows that the charge density decreases from $3\kappa E_0/4\pi$ at the tip of the boss to zero along the line of intersection of the hemisphere and the plane, and then increases to $\kappa E_0/4\pi$ at a great distance from the boss. This problem illustrates the general rule that the charge density and therefore the adjacent field is greatest where the surface of a conductor is most convex and least where the surface is most concave.

Dielectric Sphere in Uniform Field.—We will now replace the conducting sphere of Fig. 50 by an uncharged dielectric sphere of specific inductive capacity κ and suppose the region outside the sphere to be empty space. In this case the surface of the sphere need no longer be an equipotential surface, but instead it is necessary that the normal components of the electric displacement and the tangential components of the electric intensity should be the same on both sides of the surface, as was proved in article 14.

As in the problem of the conducting sphere we will take

$$V_o = A_o r \cos\theta + \frac{B_o \cos\theta}{r^2}$$

for the potential function outside the sphere. To satisfy the boundary condition at infinity it is clear that $A_o = -E_0$ as before. The components of the electric intensity are

$$(E_r)_o = -\frac{\partial V_o}{\partial r} = E_0 \cos\theta + \frac{2B_o \cos\theta}{r^3},$$

$$(E_\theta)_o = -\frac{\partial V_o}{r\partial\theta} = -E_0 \sin\theta + \frac{B_o \sin\theta}{r^3}.$$

We have supposed that there is no volume distribution of free charge inside the sphere. Therefore the potential must be a solution of Laplace's equation. We will try a function

of the same form as that used outside:

$$V_i = A_i r \cos\theta + \frac{B_i \cos\theta}{r^2}.$$

Evidently B_i must be zero, for otherwise the potential would become infinite at the origin. At the surface of the sphere V_i must equal V_o. Consequently, if a is the radius of the sphere,

$$A_i a \cos\theta = -E_0 a \cos\theta + \frac{B_o \cos\theta}{a^2}. \qquad (28\text{–}7)$$

The components of electric intensity inside the sphere are

$$(E_r)_i = -\frac{\partial V_i}{\partial r} = -A_i \cos\theta,$$

$$(E_\theta)_i = -\frac{\partial V_i}{r\partial\theta} = A_i \sin\theta.$$

The continuity of the normal component of the displacement requires that $\kappa(E_r)_i = (E_r)_o$ when $r = a$, that is,

$$-\kappa A_i \cos\theta = E_0 \cos\theta + \frac{2B_o \cos\theta}{a^3}. \qquad (28\text{–}8)$$

We need not write down the relation expressing the continuity of the tangential component of the electric intensity, for this condition is already satisfied by the equation (28–7) for the continuity of the potential.

Solving (28–7) and (28–8) for A_i and B_o we find

$$A_i = -\frac{3}{\kappa+2}E_0, \qquad B_o = \frac{\kappa-1}{\kappa+2}a^3 E_0.$$

Consequently the potential outside the sphere is

$$V_o = -\left(1 - \frac{\kappa-1}{\kappa+2}\frac{a^3}{r^3}\right)E_0 r\cos\theta, \qquad (28\text{–}9)$$

and inside,

$$V_i = -\frac{3}{\kappa+2}E_0 r\cos\theta = -\frac{3}{\kappa+2}E_0 x. \qquad (28\text{–}10)$$

Evidently the lines of force inside the sphere are parallel to

the X axis, the electric intensity having the constant value

$$(E_x)_i = -\frac{\partial V_i}{\partial x} = \frac{3}{\kappa + 2} E_0,$$

and the electric displacement being

$$(D_x)_i = \kappa(E_x)_i = \frac{3\kappa}{\kappa + 2} E_0 = \frac{3}{1 + \dfrac{2}{\kappa}} E_0.$$

The number of lines of displacement per unit cross-section inside the dielectric is to the number at a great distance in the ratio of 3 to $1 + 2/\kappa$. As κ is always greater than unity, the

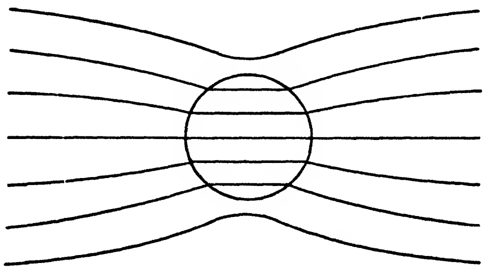


FIG. 52

lines of displacement are crowded together inside the dielectric as illustrated in Fig. 52. Note that if κ is made infinite (28–9) and (28–10) go over into the corresponding expressions for a conducting sphere. In electrostatic problems a conductor may be considered as a dielectric of infinite specific inductive capacity.

Problem 28a. A spherical conductor with charge Q surrounded by a dielectric of specific inductive capacity κ is subject to a uniform external field E_0. Find the stress on the surface and the total force on the sphere.

$$\text{Ans. } S = \frac{\kappa}{8\pi}\left\{9E_0^2\cos^2\theta + 6\frac{E_0 Q}{\kappa a^2}\cos\theta + \frac{Q^2}{\kappa^2 a^4}\right\}, \quad F = QE_0.$$

Problem 28b. Find the polarization in the dielectric sphere of this article, and the charge due to polarization per unit area of the surface.

$$\text{Ans. } P = \frac{3}{4\pi}\frac{\kappa - 1}{\kappa + 2}E_0, \quad \sigma_P = P\cos\theta.$$

Problem 28c. An isotropic dielectric of infinite extent is in a uniform field E_0. A spherical cavity of radius a is cut out of the dielectric. Find the potential V_o in the dielectric and V_i in the cavity, the polarization charge σ_P on the walls of the cavity and the field E_i inside the cavity.

$$\text{Ans. } V_o = -\left(1 + \frac{\kappa - 1}{2\kappa + 1}\frac{a^3}{r^3}\right) E_0 r \cos\theta,$$

$$V_i = -\frac{3\kappa}{2\kappa + 1} E_0 r \cos\theta,$$

$$\sigma_P = -\frac{3}{4\pi}\frac{\kappa - 1}{2\kappa + 1} E_0 \cos\theta,$$

$$E_i = \frac{3\kappa}{2\kappa + 1} E_0.$$

Problem 28d. Two point charges q and q' are placed in small spherical cavities in a rigid isotropic dielectric of infinite extent at a distance r apart. The charges do not touch the walls of the cavities. Find the force between them and compare with (16–10).

$$\text{Ans. } F = \frac{3}{2\kappa+1}\frac{qq'}{r^2}.$$

29. Conducting Cylinder in Uniform Field. — Consider an infinitely long uncharged conducting cylinder of circular cross-section placed in a uniform electric field E_0 with its axis at right angles to the lines of force. Denote the radius of the cylinder by a, and take the X axis in the direction of the field, as in Fig. 50, and the Z axis along the axis of the cylinder. It is clear from symmetry that the potential is not a function of z, and therefore the problem is one in two dimensions for the solution of which polar coordinates are indicated. The boundary conditions are evidently

$$V = -E_0 r \cos\theta \quad \text{for} \quad r = \infty,$$
$$V = 0 \quad \text{for} \quad r = a,$$

provided we take the potential of the conductor to be zero. Referring to Table II it is clear that the potential function must be of the form

$$V = Ar\cos\theta + \frac{B\cos\theta}{r},$$

the condition at infinity being satisfied by taking $A = -E_0$,

and the condition at the surface of the cylinder being satisfied by making $B = E_0a^2$. Hence the potential outside the cylinder is

$$V = -\left(1 - \frac{a^2}{r^2}\right) E_0 r \cos\theta, \qquad (29\text{-}1)$$

and the components of electric intensity in the directions of increasing r and θ are

$$E_r = -\frac{\partial V}{\partial r} = \left(1 + \frac{a^2}{r^2}\right) E_0 \cos\theta, \qquad (29\text{-}2)$$

$$E_\theta = -\frac{\partial V}{r\partial\theta} = -\left(1 - \frac{a^2}{r^2}\right) E_0 \sin\theta. \qquad (29\text{-}3)$$

The charge per unit area of the conducting surface is

$$\sigma = -\frac{\kappa}{4\pi}\left(\frac{\partial V}{\partial r}\right)_{r=a} = \frac{2\kappa E_0 \cos\theta}{4\pi}. \qquad (29\text{-}4)$$

If the cylinder, instead of being uncharged, has a charge λ per unit length, we must add to the potential function a term involving the cylindrical harmonic $\log r$ of the form $-\dfrac{2\lambda}{\kappa}\log\dfrac{r}{a}$, getting

$$V = -\left(1 - \frac{a^2}{r^2}\right) E_0 r \cos\theta - \frac{2\lambda}{\kappa}\log\frac{r}{a}. \qquad (29\text{-}5)$$

This term increases the radial component of the electric intensity by $2\lambda/\kappa r$ and the charge per unit area by $\lambda/2\pi a$.

Problem 29a. Discuss the field of a charged conducting plane of infinite extent having a cylindrical ridge of semi-circular cross-section, finding the charge per unit area on the surface of the plane.

$$\text{Ans. } \sigma = \left(1 - \frac{a^2}{r^2}\right)\frac{\kappa E_0}{4\pi}.$$

Problem 29b. A long cylindrical rod of radius a and specific inductive capacity κ is placed in a uniform electric field E_0 with its axis at right angles to the lines of force. Find the potential outside and inside the rod.

$$\text{Ans. } V_o = -\left(1 - \frac{\kappa - 1}{\kappa + 1}\frac{a^2}{r^2}\right) E_0 r \cos\theta,$$

$$V_i = -\frac{2}{\kappa + 1} E_0 r \cos\theta.$$

– *Problem 29c.* An isotropic dielectric of infinite extent lies in a uniform electric field E_0. The dielectric contains a cylindrical cavity of radius a with its axis at right angles to the field. Find the electric intensity in the cavity. Ans. $\dfrac{2\kappa}{\kappa + 1} E_0$.

Problem 29d. Find the stress on a cylindrical conductor with charge λ per unit length surrounded by a dielectric and subject to a uniform field E_0 at right angles to its axis. What is the force per unit length of the cylinder?

$$\text{Ans.}\quad S = \frac{\kappa}{8\pi}\left\{4E_0^2\cos^2\theta + 8\frac{\lambda}{\kappa a}E_0\cos\theta + \frac{4\lambda^2}{\kappa^2 a^2}\right\},\quad F = \lambda E_0.$$

30. Electrical Images. — Consider the electric field produced by two point charges. If we replace an equipotential surface surrounding one of the charges by a conducting surface and then transfer the charge to this surface the field between the other charge and the surface remains unaltered, although that between the first charge and the surface is wiped out. For the potential in the original field satisfies Laplace's equation, and the transfer of the electricity on one of the point charges to a conducting surface coinciding with one of the equipotential surfaces of the field merely replaces the original boundary conditions by equivalent boundary conditions in so far as the portion of the field remaining is concerned. The method of solving electrostatic problems by the use of electrical images consists in its simplest form in placing two point charges in such positions that one of the equipotential surfaces of the field produced coincides with the surface of a conductor which it is desired to place in the field. Then the charge on one side of the conducting surface is transferred to it, the field on the other side remaining unaltered. When this method is feasible it enables us to determine the field produced by a point charge and a conducting surface from the simple investigation of the field due to two point charges. In such a case the point charge which is transferred to the conducting surface is said to be the *image* of the other point charge.

The method may be extended so as to find the field produced by two conductors neither of which constitutes a point charge, such as the field due to two charged conducting spheres. In

problems in which the potential is not a function of one of the rectangular coordinates the point charges referred to above are replaced by line charges.

Point and Plane.—To find the field due to a point charge q (Fig. 53) and a conducting plane AB at a distance $d/2$ from it, we consider the field due to q and a point charge $-q$ placed on the perpendicular dropped from q to the plane as far to the left

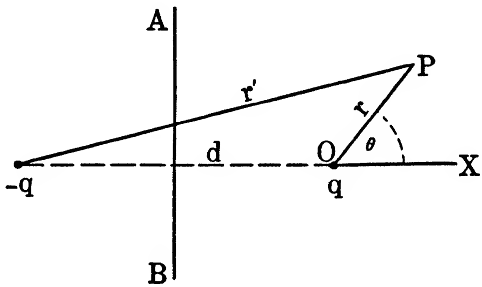


FIG. 53

of the plane as q is to the right. The potential at P due to the two point charges is evidently

$$V = \frac{q}{\kappa r} - \frac{q}{\kappa r'}$$

$$= \frac{q}{\kappa r} - \frac{q}{\kappa\sqrt{r^2 + 2rd\cos\theta + d^2}}, \qquad (30\text{–}1)$$

where κ is the specific inductive capacity of the medium surrounding the charges.

Evidently the median plane AB is an equipotential surface of zero potential. So if we transfer the charge on the image $-q$ to an earthed conducting plane placed so as to coincide with AB the field to the right of AB remains unaltered while that to the left is wiped out. The potential function (30–1) still applies to the region to the right of AB and the components of electric intensity are

$$E_r = -\frac{\partial V}{\partial r} = \frac{q}{\kappa r^2} - \frac{q(r + d\cos\theta)}{\kappa(r^2 + 2rd\cos\theta + d^2)^{3/2}}, \qquad (30\text{–}2)$$

$$E_\theta = -\frac{\partial V}{r\partial\theta} = \frac{qd\sin\theta}{\kappa(r^2 + 2rd\cos\theta + d^2)^{3/2}}. \qquad (30\text{-}3)$$

The density of the charge induced on the earthed conducting plane by q is

$$\sigma = -\frac{\kappa}{4\pi}\left(\frac{\partial V}{\partial x}\right)_{x=-d/2} = \frac{\kappa}{4\pi}(E_x)_{x=-d/2}$$

$$= \frac{\kappa}{4\pi}(E_r\cos\theta - E_\theta\sin\theta)_{r=r'} = -\frac{qd}{4\pi r^3}. \qquad (30\text{-}4)$$

As the field in the neighborhood of q is not changed by the substitution of the conducting plane for the image $-q$ it follows from the law of action and reaction that the total force on the conducting plane due to the attraction of q is the same as that which would be experienced by the image $-q$, that is,

$$F = -\frac{q^2}{\kappa d^2}. \qquad (30\text{-}5)$$

Point and Sphere.—In this case it is required to find the magnitude and position of the point charge $-q'$ (Fig. 54) so

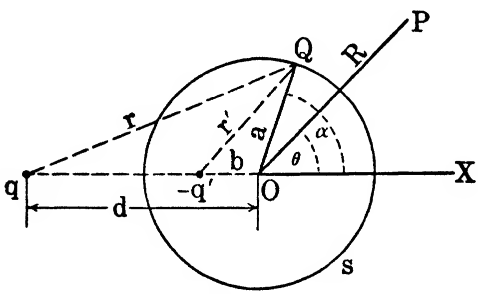


FIG. 54

that the spherical surface s of radius a shall be an equipotential in the field produced by the charges q and $-q'$. If the distances of q and $-q'$ from the center O of the sphere are denoted by d and b respectively, the potential at a point Q on the surface of the sphere is

$$V = \frac{q}{\kappa r} - \frac{q'}{\kappa r'} = \frac{q}{\kappa\sqrt{d^2 + 2da\cos\alpha + a^2}} - \frac{q'}{\kappa\sqrt{a^2 + 2ab\cos\alpha + b^2}}.$$

If we make $b = a^2/d$,

$$V = \frac{q - \dfrac{d}{a}q'}{\kappa\sqrt{d^2 + 2da\cos\alpha + a^2}},$$

and if now we give q' the magnitude $(a/d)q$ the potential vanishes for all values of α. Therefore the sphere is a surface of zero potential. Now we replace the sphere by an earthed conducting surface and transfer the charge $-q'$ to it. This annuls the field inside the sphere leaving the field outside unaltered. Therefore the potential at any exterior point P due to the point charge q and the earthed sphere is

$$V = \frac{q}{\kappa\sqrt{R^2 + 2Rd\cos\theta + d^2}} - \frac{\left(\dfrac{a}{d}\right)q}{\kappa\sqrt{R^2 + 2R\dfrac{a^2}{d}\cos\theta + \dfrac{a^4}{d^2}}}. \qquad (30\text{–}6)$$

The induced charge per unit area on the surface of the sphere is

$$\sigma = -\frac{\kappa}{4\pi}\left(\frac{\partial V}{\partial R}\right)_{R=a} = -\frac{q}{4\pi a}\frac{d^2 - a^2}{(d^2 + 2da\cos\theta + a^2)^{3/2}}. \qquad (30\text{–}7)$$

As in the previous case, the force between the charge q and the sphere is equal to that between q and its image $-q'$ in the sphere.

In the case we have been considering the potential of the sphere has been made zero. If we wish the sphere to have a potential other than zero we can add a uniformly distributed charge Q to the surface of the sphere, for the effect of such a charge is merely to increase the potential of every point on the surface by $Q/\kappa a$ and therefore the surface remains equipotential. The potential at an exterior point P will then exceed (30–6) by $Q/\kappa R$.

If instead of transferring $-q'$ to the surface of the conducting spherical shell we leave it in its original position and transfer q to the spherical surface, the field inside the sphere is unaltered and therefore the potential at any *interior* point is given by (30–6) to within an additive constant. The field outside the sphere,

however, is not annulled in this case. For only a portion q' of the charge q resides on the inner surface of the sphere, all the lines of force originating on it ending on $-q'$. The remainder of the charge q, namely,

$$q - q' = \frac{d - a}{d} q,$$

spreads itself uniformly over the outer surface of the sphere, giving rise to the potential

$$\frac{d - a}{\kappa R d} q$$

at exterior points. For the screening effect of the conductor prevents the field due to $-q'$ and the charge q' on the inner surface of the shell from penetrating to the outside.

Problem 30a. What is the effect of adding a uniformly distributed charge to the conducting plane of Fig. 53?

Problem 30b. Find the potential of an uncharged spherical conductor of radius a whose center is at a distance d from a point charge q. Ans. $\frac{q}{,}$.

31. Sphere and Plane. — We shall apply the method of images to find the capacity of a charged spherical conductor of radius a (Fig. 55) at a distance $d/2$ from an earthed conducting plane AB of infinite extent, a being supposed small compared to d.

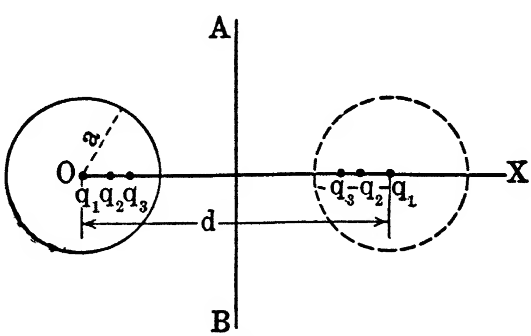


FIG. 55

In this case we must make use of an infinite series of images. First suppose a charge q_1 placed at the center O of the sphere. This charge gives rise to a field of which the surface of the sphere is an equipotential, but the surface of the plane is not equipotential. Adding the image $-q_1$ of q_1 in the plane, AB becomes equipotential, but the sphere is no longer so. Next add the image q_2 of $-q_1$ in the sphere; this restores the equipotential character of the spherical surface but disturbs that of the plane. Addition of the image $-q_2$ of q_2 in the plane makes the plane an equipotential surface again but disturbs the potential over the surface of the sphere. Each one of the added fields, however, is smaller than the preceding one, so continuation of the process brings us ever nearer to the desired state in which both the plane and the spherical surface are equipotential surfaces. Evidently the total charge on the sphere is the sum of the series $q_1 + q_2 + q_3 + \cdots$ and the charge on the plane is equal but opposite in sign.

The distance of $-q_1$ from the center of the sphere is d, so it follows from the previous article that q_2 is a charge $(a/d)q_1$ at a distance a^2/d from O. Therefore $-q_2$ is at a distance $d\left(1 - \frac{a^2}{d^2}\right)$ from O and q_3 is a charge

$$\frac{\frac{a}{d}}{1 - \frac{a^2}{d^2}} q_2 = \frac{\frac{a^2}{d^2}}{1 - \frac{a^2}{d^2}} q_1$$

at a distance

$$\frac{\frac{a^2}{d}}{1 - \frac{a^2}{d^2}}$$

from O. If we put m for the ratio a/d we have then the following point charges at the distances specified from the center of the

sphere:

Charge	*Distance from O*
$q_1,$	$0,$
$mq_1,$	$ma,$
$\dfrac{m^2}{1-m^2}q_1,$	$\dfrac{m}{1-m^2}a,$
$\dfrac{m^3}{(1-m^2)\left(1-\dfrac{m^2}{1-m^2}\right)}q_1,$	$\dfrac{m}{1-\dfrac{m^2}{1-m^2}}a,$
$\dfrac{m^4}{(1-m^2)\left(1-\dfrac{m^2}{1-m^2}\right)\left(1-\dfrac{m^2}{1-\dfrac{m^2}{1-m^2}}\right)}q_1,$	$\dfrac{m}{1-\dfrac{m^2}{1-\dfrac{m^2}{1-m^2}}}a,$
. . . .	. . . .

The total charge on the sphere is

$$Q = q_1\left\{1 + m + \frac{m^2}{1-m^2} + \frac{m^3}{(1-m^2)\left(1-\dfrac{m^2}{1-m^2}\right)}\right.$$

$$\left. + \frac{m^4}{(1-m^2)\left(1-\dfrac{m^2}{1-m^2}\right)\left(1-\dfrac{m^2}{1-\dfrac{m^2}{1-m^2}}\right)} + \cdots\right\}$$

$$= q_1\{1 + m + m^2 + m^3 + 2m^4 + 3m^5 + \cdots\}. \qquad (31\text{–}1)$$

Since q_2 is the image of $-q_1$ in the sphere, the potential of the surface of the sphere due to this pair of charges is zero, as shown in the previous article. The same is true of the pairs q_3 and $-q_2$, q_4 and $-q_3$, etc. Consequently the potential of the sphere is

$$V = \frac{q_1}{\kappa a}, \qquad (31\text{–}2)$$

due to the image charge q_1 at its center alone.

Since equal charges of opposite sign are placed symmetrically on the two sides of AB, the potential of the plane is zero. Con-

sequently the capacity of a sphere relative to an earthed plane is

$$C = \frac{Q}{V} = \kappa a\{1 + m + m^2 + m^3 + 2m^4 + 3m^5 + \cdots\}. \quad (31\text{–}3)$$

Note that the capacity of the sphere is increased by the proximity of the plane.

The analysis above provides us at once with the solution of the problem of two conducting spheres of the same radius a a distance d apart, the charge on the second sphere being equal but opposite in sign to that on the first. The negatively charged sphere is indicated by a broken circle in Fig. 55. Its potential is

$$V = -\frac{q_1}{\kappa a}. \quad (31\text{–}4)$$

The field on the right of AB is the reflection in the plane of the field on the left.

Problem 31a. Show that the coefficient of induction is the negative of the capacity (31–3).

Problem 31b. A sphere of 10 cm radius is suspended 25 cm above the ground. Calculate the capacity of the sphere relative to earth. To how many figures is the answer correct? Ans. 12.52 cm.

32. Parallel Wires. — Consider two infinitely long straight parallel wires of radius a, shown in cross-section in Fig. 56. We wish to find the field when one has a charge λ per unit length and the other an equal negative charge. Then we shall compute the capacity of the one relative to the other.

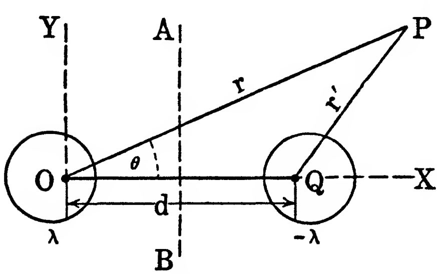


Fig. 56

Begin by considering two parallel filaments intersecting the plane of the figure at O and Q and having charges λ and $-\lambda$ per unit length respectively. Evidently the potential at a point P due to them is not a function of the coordinate z at right angles

to the plane of the figure. So it must be represented by one of the cylindrical harmonics of Table II, article 27. From symmetry it is clear that the potential due to each filament is a function only of the distance of P from the filament; therefore we have

$$V = -\frac{2\lambda}{\kappa}\log r + \frac{2\lambda}{\kappa}\log r' = \frac{\lambda}{\kappa}\log\frac{r'^2}{r^2}. \qquad (32\text{–}1)$$

This potential function makes the potential zero midway between the two filaments and satisfies the condition that the charges on the filaments shall be λ and $-\lambda$ per unit length.

The traces of the equipotential surfaces are

$$\frac{r'^2}{r^2} = \frac{r^2 - 2rd\cos\theta + d^2}{r^2} \equiv m^2,$$

where m is a constant, or

$$r^2 + 2r\cos\theta\left(\frac{d}{m^2-1}\right) = \frac{d^2}{m^2-1}.$$

Transforming to rectangular coordinates $x = r\cos\theta$, $y = r\sin\theta$,

$$\left(x + \frac{d}{m^2-1}\right)^2 + y^2 = \frac{m^2d^2}{(m^2-1)^2}.$$

This is the equation of a circle of radius $\dfrac{md}{m^2-1}$ with center a distance $\dfrac{d}{m^2-1}$ to the left of the origin. The equipotential surfaces are cylinders of circular cross-section, and we may take the two of radius a as the surfaces of the wires, transferring the charge on each filament to the surrounding cylindrical surface and annulling the field inside. The radius a of the wires is then

$$a = \frac{md}{m^2-1},$$

and the distance b between the axes of the two wires is

$$b = d + \frac{2d}{m^2-1} = \frac{m^2+1}{m^2-1}d.$$

Eliminating d between these two equations,

$$m^2 - \frac{b}{a}m + 1 = 0,$$

of which the roots are

$$m = \frac{b \pm \sqrt{b^2 - 4a^2}}{2a}.$$

Evidently the positive sign before the radical gives the root m_+ pertaining to the positively charged wire and the negative sign the root m_- pertaining to the negatively charged wire. Hence the potential of the first is

$$V = \frac{2\lambda}{\kappa}\log m_+ = \frac{2\lambda}{\kappa}\log\frac{b + \sqrt{b^2 - 4a^2}}{2a},$$

and that of the second,

$$V' = \frac{2\lambda}{\kappa}\log m_- = \frac{2\lambda}{\kappa}\log\frac{b - \sqrt{b^2 - 4a^2}}{2a}.$$

The difference of potential is

$$V - V' = \frac{2\lambda}{\kappa}\log\frac{b + \sqrt{b^2 - 4a^2}}{b - \sqrt{b^2 - 4a^2}}, \qquad (32\text{–}2)$$

and the capacity per unit length,

$$C = \frac{\lambda}{V - V'} = \frac{\kappa}{2\log\dfrac{b + \sqrt{b^2 - 4a^2}}{b - \sqrt{b^2 - 4a^2}}}. \qquad (32\text{–}3)$$

We have noted that the median plane AB is an equipotential surface of zero potential. Therefore we might have transferred the charge $-\lambda$ per unit length on the right-hand filament to a conducting plane placed so as to coincide with AB, wiping out the field to the right of the plane. This device enables us to calculate the capacity of a horizontal wire relative to earth. If $h = b/2$ is the elevation of the axis of the wire,

$$V = \frac{2\lambda}{\kappa}\log\frac{h + \sqrt{h^2 - a^2}}{a},$$

and the capacity per unit length is

$$C = \frac{\lambda}{V} = \frac{\kappa}{2\log\dfrac{h+\sqrt{h^2-a^2}}{a}}. \qquad (32\text{–}4)$$

Problem 32a. A horizontal wire of negligible radius with a charge λ per unit length is suspended at a height h above the surface of a conducting plane. Find the potential as a function of r and θ, where r is the distance from the wire and θ the angle with the downward drawn vertical, and find the charge per unit area on the surface of the plane.

$$\text{Ans. } V = \frac{\lambda}{\kappa}\log_e\left(1 - 4\frac{h}{r}\cos\theta + 4\frac{h^2}{r^2}\right), \quad \sigma = -\frac{\lambda h}{\pi r^2}.$$

Problem 32b. Find the force per unit length between the two wires of article 32. Ans. $-\dfrac{2\lambda^2}{\kappa d}$.

—*Problem 32c.* What is the capacity per unit length of a wire of 5 mm diameter (*a*) relative to a parallel wire at a distance of 13 mm, (*b*) relative to an earthed plane at the same distance? Assume $\kappa = 1$. Ans. (*a*) 0.155, (*b*) 0.214.

—*Problem 32d.* A wire of 3 mm radius is 5 mm from an earthed conducting plane. When the wire is raised to a potential of 1000 volts what charge is induced on the plane per unit length of the wire? Ans. $-0.505(10)^{-9}$ coulomb/cm.

33. Conjugate Functions. — An electrostatic problem in which the potential is not a function of one of the rectangular coordinates, say z, is called a two-dimensional problem. In such a case Laplace's equation (25–2) reduces to

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0. \qquad (33\text{–}1)$$

We have seen how to solve such a problem in the case where V is expressed in polar coordinates r and θ by the use of cylindrical harmonics. Now we shall consider a more general method involving what are known as *conjugate functions.*

Let $\mathbf{z}$ represent the complex quantity

$$\mathbf{z} \equiv x + iy,$$

where $\mathrm{i} \equiv \sqrt{-1}$. To distinguish complex from real quantities we shall use **black face** type to designate the former. A more detailed discussion of such quantities is given in article 111.

Consider any function of $\mathbf{z}$ such as $F(\mathbf{z})$, and let $g(x, y)$ and $h(x, y)$ be respectively the real and the imaginary parts of $F(\mathbf{z})$, so that

$$F(\mathbf{z}) = F(x + \mathrm{i}y) = g(x, y) + \mathrm{i}h(x, y).$$

The functions g and h are called *conjugate functions* of x and y.

Now

$$\left.\begin{aligned} \frac{\partial F}{\partial x} &= \frac{dF}{d\mathbf{z}}, & \frac{\partial^2 F}{\partial x^2} &= \frac{d^2F}{d\mathbf{z}^2}; \\ \frac{\partial F}{\partial y} &= \mathrm{i}\frac{dF}{d\mathbf{z}}, & \frac{\partial^2 F}{\partial y^2} &= -\frac{d^2F}{d\mathbf{z}^2}. \end{aligned}\right\} \tag{33-2}$$

Therefore

$$\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} = \frac{d^2F}{d\mathbf{z}^2} - \frac{d^2F}{d\mathbf{z}^2} = 0. \tag{33-3}$$

So any function of the complex variable $x + \mathrm{i}y$ satisfies Laplace's equation (33-1) for two dimensions. Consequently the real part $g(x, y)$ of $F(\mathbf{z})$ must satisfy Laplace's equation and the same is true for the imaginary part $h(x, y)$. Hence either the real part of $F(\mathbf{z})$ or the imaginary part may be taken as the potential function in a possible electrostatic field.

From (33-2) we have

$$\frac{\partial F}{\partial y} = \mathrm{i}\frac{\partial F}{\partial x},$$

and therefore

$$\frac{\partial g}{\partial y} + \mathrm{i}\frac{\partial h}{\partial y} = \mathrm{i}\left(\frac{\partial g}{\partial x} + \mathrm{i}\frac{\partial h}{\partial x}\right).$$

So, as g and h are real functions,

$$\frac{\partial g}{\partial y} = -\frac{\partial h}{\partial x}, \qquad \frac{\partial h}{\partial y} = \frac{\partial g}{\partial x},$$

and

$$\frac{\dfrac{\partial g}{\partial x}}{\dfrac{\partial g}{\partial y}} = -\frac{\dfrac{\partial h}{\partial y}}{\dfrac{\partial h}{\partial x}}. \tag{33-4}$$

Now consider the two families of curves

$$g(x, y) = \text{constant}, \qquad h(x, y) = \text{constant}.$$

The slope of the tangent to a curve of the first family is

$$\left(\frac{dy}{dx}\right)_g = -\frac{\dfrac{\partial g}{\partial x}}{\dfrac{\partial g}{\partial y}},$$

and that to a curve of the second family is

$$\left(\frac{dy}{dx}\right)_h = -\frac{\dfrac{\partial h}{\partial x}}{\dfrac{\partial h}{\partial y}}.$$

But (33–4) states that the one slope is the negative reciprocal of the other, that is, that the two families of curves intersect orthogonally. So if one family represents the traces of equipotential surfaces on the XY plane the other represents lines of force.

Problem 33a. Show that the function

$$F(z) = (x + iy)^n = (re^{i\theta})^n$$

leads to the cylindrical harmonics of article 27.

Problem 33b. Show, by finding the function conjugate to $\log r$, that the equation of the lines of force between the two wires of article 32 is

$$\theta' - \theta = \text{constant},$$

where θ' is the angle which r' makes with the X axis.

34. Examples of Conjugate Functions. — We shall consider a few examples of the use of conjugate functions which are of practical interest.

Field Due to Two Conducting Planes Intersecting at Right Angles.—The function of the complex variable $\mathbf{z}$ needed for the solution of this problem is

$$F(\mathbf{z}) = A(x + iy)^2 = A(x^2 - y^2) + 2iAxy, \qquad (34\text{–}1)$$

where we choose for the potential function the imaginary part,

$$V = 2Axy. \tag{34-2}$$

The equipotential surfaces are the equilateral hyperbolic cylinders whose traces are shown by full lines in Fig. 57. The conducting planes, which coincide with the XZ and YZ coordinate planes, are at zero potential. The equations of the lines of force are the hyperbolas,

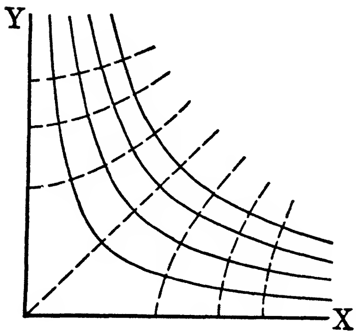


FIG. 57

$$x^2 - y^2 = \text{constant}, \tag{34-3}$$

shown by broken lines in the figure.

The components of the electric intensity are

$$\left.\begin{aligned} E_x &= -\frac{\partial V}{\partial x} = -2Ay, \\ E_y &= -\frac{\partial V}{\partial y} = -2Ax, \end{aligned}\right\} \tag{34-4}$$

and the charge per unit area on the XZ plane is

$$\sigma = -\frac{\kappa}{4\pi}\left(\frac{\partial V}{\partial y}\right)_{y=0} = -\frac{\kappa A}{2\pi}x, \tag{34-5}$$

with a similar expression for that on the YZ plane. The density of charge, then, is zero along the line of intersection and increases linearly with the distance from this line. If A is positive the charge on the conducting planes is negative since the potential increases as x and y are made greater. The lines of force are directed toward the planes and terminate on negative charges on their surfaces.

Field at the Edge of a Conducting Plane.—For this case the appropriate function is

$$F(\mathbf{z}) = \{A(x + iy)\}^{1/2} = g(x, y) + ih(x, y). \tag{34-6}$$

Therefore

$$Ay =$$

and, choosing $h(x, y)$ for the potential function V, we get, on eliminating $g(x, y)$,

$$A^2y^2 = 4V^2(Ax + V^2). \tag{34-7}$$

The equipotential surfaces are parabolic cylinders as illustrated in Fig. 58, the positive half of the X axis being the trace of the conducting plane, which is at zero potential.

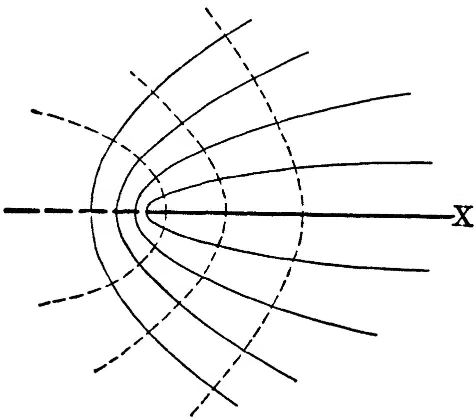


FIG. 58

To get the charge per unit area differentiate (34–7) partially with respect to y. Then

$$\frac{\partial V}{\partial y} = \frac{A^2y}{4V(Ax + 2V^2)} = \frac{A\sqrt{Ax + V^2}}{2(Ax + 2V^2)}$$

on eliminating y by means of (34–7). Putting $V = 0$ is equivalent to making y zero, so

$$\sigma = -\frac{\kappa}{4\pi}\left(\frac{\partial V}{\partial y}\right)_{y=0} = -\frac{\kappa}{8\pi}\sqrt{\frac{A}{x}}. \tag{34-8}$$

This gives the charge on one side of the plane only. The charge density is infinite at the edge where the convexity of the surface becomes infinitely great.

CHAPTER IV

MAGNETOSTATICS

35. Attraction and Repulsion. — The use of a natural magnet or *lodestone* to indicate direction at sea appears to have been unknown to the ancients, but was certainly familiar to European navigators at the time of the Crusades. As early as the thirteenth century Peregrinus found that, if an iron needle is suspended at different places above the surface of a lodestone and lines drawn so as to have everywhere the direction assumed by the needle, these lines converge at two points on opposite sides of the stone. These points he named *poles*.

Little further progress was made until 1600 when Gilbert noted that there are two kinds of poles, *positive* or *north-seeking poles*, and *negative* or *south-seeking poles*, which are characterized by the fact that like poles repel and unlike poles attract each other. He also explained the action of a freely suspended magnet in assuming a north and south orientation as due to the presence of poles in the earth roughly coincident with the geographical poles. As the north pole of a magnet turns toward the north, it must be attracted by a south magnetic pole in the neighborhood of the north geographical pole and repelled by a north magnetic pole near the south geographical pole. Acutally the earth's magnetic field is better represented by supposing it to be due to a comparatively short magnet placed at the center of the earth along a line inclined 17° to the geographical axis.

The early discoveries in magnetism, then, may be summed up in the two qualitative laws:

(I) *Like poles repel, unlike poles attract.*

(II) *The force between two poles decreases as the distance between them is increased.*

These laws are quite analogous to the corresponding laws for electric charges given in article 1. There are, however, certain

fundamental differences between electric charges and magnetic poles. In the first place, every magnet is found to have equal quantities of north and south magnetism. If, for instance, we place a magnet on a cork floating on water, the earth's field, while it may turn the magnet, gives rise to no motion of translation. This shows that the northward force on the north pole is just balanced by the southward force on the south pole, and therefore that the strength of the two poles is the same. Moreover, if a needle, magnetized so as to have a north pole at one end and a south pole at the other, is cut in two, we find that we have two complete magnets, each of which has equal north and south poles. Therefore it is impossible to isolate north and south poles on different bodies, as can be done in the case of electric charges of opposite sign. If we wish to deal experimentally with a north pole alone, the best we can do is to employ a needle so long that the force exerted by the south pole at the remote end is negligible in the neighborhood of the north pole.

A second difference between magnetostatic and electrostatic phenomena lies in the absence of any such substance as a conductor of magnetism. No substance is known in which magnetic poles move under the influence of magnetic forces with the freedom that electric charges move in a conductor when subject to electric forces. The poles in a piece of very soft iron—which is the nearest thing to a conductor of magnetism—never move so far under the action of impressed magnetic forces as to annul the field in the interior of the body.

On the other hand there are many similarities between magnetic and electric phenomena. For instance, an unmagnetized piece of iron, when placed near a pole of a strong magnet, becomes magnetized by induction in much the same way that a conductor is charged inductively when placed in an electric field, although the analogy with a dielectric is closer.

Just as in electrostatics we are concerned only with charges at rest in the observer's inertial system, so in magnetostatics we limit ourselves to a consideration of magnets at rest relative to the observer's reference frame. It is important to note that the

phenomena of electrostatics and of magnetostatics are quite distinct; no force is exerted by an electric charge on a magnetic pole when both are at rest relative to the observer, and *vice versa.*

36. Molecular Theory of Magnetism. — The fact that two magnets are produced by cutting a magnet in two suggests that if the process could be carried far enough the individual molecules of a magnetic substance would be found to be themselves tiny magnets. The molecular theory of magnetism, proposed by Weber and elaborated by Ewing, accounts for many of the properties of magnetic substances. On this theory each molecule is supposed to contain one or more small permanent magnets which are free to turn inside the molecule. In an unmagnetized iron bar the molecular magnets are arranged chaotically, so that in any volume element large enough to be investigated experimentally as many have their axes pointing in one direction as in the opposite. Therefore the bar as a whole gives rise to no magnetic forces at outside points. If, now, the north pole N (Fig. 59) of a strong magnet is brought near to one end of the

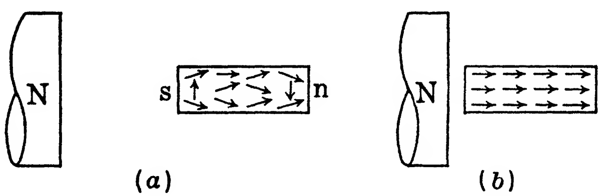


FIG. 59

bar, the molecular magnets, represented by arrows with heads at the north ends, experience a torque due to the attraction of N for the south pole of each molecule and its repulsion on the north pole. This torque tends to line up the molecular magnets parallel to the axis of the bar. Each individual magnet, however, was in equilibrium under the magnetic torques exerted by its neighbors in the original chaotic condition. Therefore the torque due to N is opposed by these internal torques, and so long as the bar is at a distance from N the molecular magnets are only

partially lined up as in (*a*). Nevertheless there is a preponderance of south poles at the end *s* of the bar and of north poles at the end *n*, while in the interior each north pole is compensated by a neighboring south pole. Thus the bar has become *magnetized by induction* with a south pole on the end adjacent to *N* and a north pole on the farther end.

Further approach of *N* toward the bar increases the torques exerted by the former on the molecular magnets, and they swing more and more into line. At a certain distance a condition approaching instability is reached, the internal torques becoming very small, and the molecular magnets turning very rapidly as the distance from *N* is decreased. If the strength of the pole *N* is great enough, a point is finally reached where the molecular magnets are all in line, as shown in (*b*). In this state the iron bar is said to be *saturated.* Since the elementary magnets of which it is composed are completely lined up, further increase of the magnetizing force cannot increase the strength of the poles on its ends.

If, now, the pole *N* is removed the molecular magnets are no longer subject to a torque due to external causes. They do not, however, resume their original orientations, for now each is acted on by a different set of internal torques from that existing in the original unmagnetized state. Instead the bar retains a certain fraction of the induced magnetization, which, however, may be partially or completely destroyed by shaking up the molecules, either by tapping the bar with a hammer or by raising its temperature.

A graph showing how the degree of magnetization of the bar varies with the magnetizing force exerted by the pole *N* is given in Fig. 60. Starting with the bar in the unmagnetized condition represented by the point *O*, we pass through the state *a* where the magnetization increases very rapidly with increase of the magnetizing force to saturation at *b*. Decreasing the magnetizing force to zero the bar is left with a degree of magnetization represented by the ordinate *Oc*. This ordinate is a measure of the *retentiveness* of the bar. Reversing the direction of the

magnetizing force and increasing its magnitude brings us to the state *d* in which the bar has completely lost its magnetization, and finally to the condition *e* of saturation in the opposite sense to that previously existing. The magnetizing force *Od* necessary to completely demagnetize the bar is often referred to as the *coercive force*. Continuation of the process takes us through the points *f* and *g* to *b* again. The lagging of the magnetization behind the magnetizing force illustrated by the curve is known as *hysteresis* and *bcefb* is called a *hysteresis loop*.

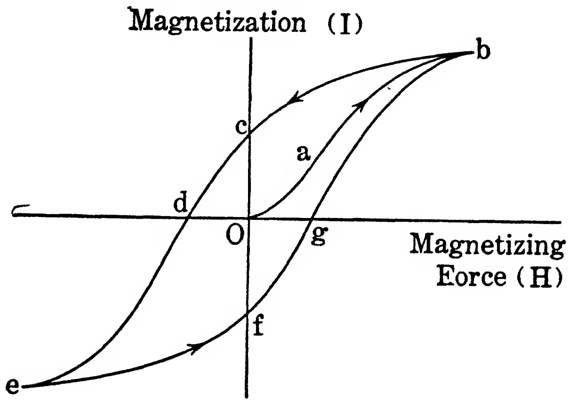


FIG. 60

The preceding discussion applies to those substances known as *ferromagnetic*, that is, iron, nickel, cobalt and certain alloys. Another class of magnetic substances, known as *paramagnetic*, exhibits similar magnetic properties to a much smaller degree but does not show the phenomena of saturation and of hysteresis in such fields as are available in the laboratory. It is probable, however, that such substances also would become saturated if subjected to sufficiently large magnetizing forces. Finally certain substances, when acted on by a magnetizing force, become weakly magnetized in the opposite sense to ferromagnetic or paramagnetic substances. These are said to be *diamagnetic*. Thus a bar of diamagnetic material, placed as in Fig. 59, would acquire a north pole at the end *s* and a south pole at *n*. While the

molecular theory of magnetism accounts satisfactorily for the general features of ferromagnetism and paramagnetism, diamagnetism is due to an entirely different cause. Its explanation will be taken up in Chapter VIII, article 83, where it will be shown that probably all substances are diamagnetic, the diamagnetism of a paramagnetic medium being masked by the very much more intense paramagnetism.

37. Coulomb's Law. — As the two kinds of poles have opposite properties, it is convenient to speak of a north pole as a *positive magnetic charge* and of a south pole as a *negative magnetic charge*. Consider two magnetic charges m and m' located at points a distance r apart. Doubling the charge m doubles the force on m', for each half of the doubled charge exerts the same force on m' as that exerted by the original charge. Therefore the force between the two poles is proportional to m. Similarly it is proportional to m'. Consequently *the force between two poles is proportional to the product of their pole-strengths or magnetic charges.*

When the distance between two point poles is changed, it is found that *the force exerted by either on the other is inversely proportional to the square of the distance between them.* This law was stated by Michell in 1750, but it is usually attributed to Coulomb who put it on a firm foundation by his experiments with the torsion balance in 1785. These experiments are the same in principle as those performed by the same investigator in determining how the force between electric charges varies with the distance. A long magnetic needle ns (Fig. 61) is suspended at its center of mass by the torsion fibre F of the torsion balance. When the north pole N of a second long magnet is brought up to a pre-

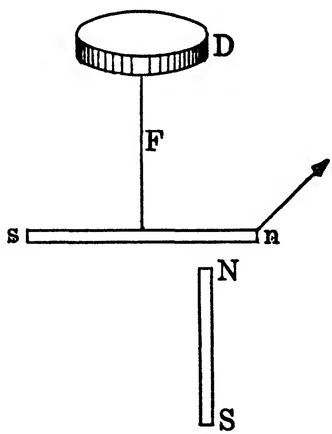


Fig. 61

determined point in front of the north pole n of the suspended magnet, the latter experiences a torque and turns about the fibre as axis. Since the fibre is rigidly attached at both ends, the rotation of ns causes it to twist. Next the drum D is turned far enough in the opposite sense to bring the suspended magnet back to its original orientation. The angle through which the drum has to be turned is a measure of the force between N and n in their final positions. By varying the distance between the two poles the relation stated above is found.

In performing the experiment it is necessary to use long magnets so that the influence of the south poles s, S will be inappreciable. Moreover it is important not to bring the two north poles too close together, as each has a tendency to weaken the other by induction. A more precise method of verifying the law of force, devised later by Gauss, will be described in article 47.

As the force F between two point poles m and m' varies directly with the product of the pole-strengths and inversely with the square of the distance r between them, we may write

$$F = \frac{mm'}{r^2}, \tag{37–1}$$

a positive force indicating repulsion and a negative force attraction. As we shall see later, the presence of a magnetic medium in the region between the poles changes the force between them. So equation (37–1) applies exactly only to poles *in vacuo*, although it may be used for poles in air in all but very precise measurements.

Giving all the quantities in (37–1) the magnitude unity, we see that this equation defines *the c.g.s. unit of pole-strength or magnetic charge as that pole which repels a like equal pole placed at a distance of 1 cm in vacuo with a force of 1 dyne.* Putting 1 gm cm/sec^2 for F and 1 cm for r we find 1 gm$^{1/2}$ cm$^{3/2}$/sec for the c.g.s. unit of magnetic charge. This unit, and all other units derived from the fundamental equation (37–1), are known as *electromagnetic units* and are designated by the abbreviation e.m.u. Although the electromagnetic unit of pole-strength has the same

physical dimensions as the electrostatic unit of charge, the student must remember that magnetostatic phenomena are quite distinct from those of electrostatics, and that it is not until we study currents that we can establish a relation between these two units.

38. Field Strength and Potential. — In order to explore a *magnetic field*—that is, a region in which magnetic forces are acting—we may suppose a unit north pole of very small dimensions to be carried around the field, all the poles producing the field being kept fixed and of constant strength. The force experienced by the test pole when at rest relative to the observer at any point in the field is known as the *field strength*, *magnetic intensity*, or *magnetic force* H at that point. Since magnetic intensity is a force it is a vector quantity, its magnitude being measured in *dynes per unit pole*. The electromagnetic unit of H is also called the *gauss*.

The force F on a pole of strength m placed at a point in a magnetic field where the magnetic intensity is H is given by the equation,

$$F = mH. \tag{38–1}$$

If m is positive F has the direction of H, otherwise the opposite direction. The magnetic intensity at a distance r from a point pole of strength m, which is obtained from (37–1) by making m' unity, is

$$H = \frac{m}{r^2} \tag{38–2}$$

in the direction of the radius vector drawn from m.

The potential V at a point in a magnetic field is the work necessary to bring a unit positive pole from outside the field up to the point in question, all the poles producing the field being kept fixed and constant in strength during the process. It is equal to the potential energy of a unit pole placed at the point in question and is measured in *ergs per unit pole*.

To calculate the potential produced at P (Fig. 62) by a point pole m placed at O we compute the work done in bringing a unit positive pole from an infinitely distant point Q to P along the path QP. At Q_2 the force is

$$H = \frac{m}{r^2},$$

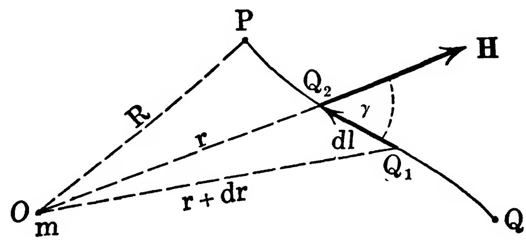


Fig. 62

and the work done in carrying the unit pole from Q_1 to Q_2 is

$$dV = \frac{m}{r^2}\, dl \cos \gamma.$$

As
$$dl \cos \gamma = -\, dr,$$

$$dV = -\frac{m}{r^2}\, dr$$

and, integrating from ∞ to R, the potential at P is found to be

$$V = -\, m \int_\infty^R \frac{dr}{r^2} = \frac{m}{R}. \qquad (38\text{–}3)$$

In the case of a number of point poles of strengths m_1, m_2, m_3, $\cdots$ at distances r_1, r_2, r_3, $\cdots$ from P the potential is

$$V = \frac{m_1}{r_1} + \frac{m_2}{r_2} + \frac{m_3}{r_3} + \cdots = \Sigma \frac{m}{r}, \qquad (38\text{–}4)$$

and in the case of a continuous distribution of magnetic charge consisting of ρ units per unit volume inside a volume τ and σ

units per unit area on a surface s,

$$V = \int_\tau \frac{\rho d\tau}{r} + \int_s \frac{\sigma ds}{r}. \tag{38–5}$$

Evidently the potential in a magnetostatic field is a function only of the coordinates of the point P at which it is to be evaluated, and is independent of the path along which the unit pole is carried to P.

It follows from the definition of potential that the difference in potential of two points P and Q is measured by the work necessary to carry a unit positive pole from the one to the other, all paths between the two points being equivalent. As the change in potential in going around a closed curve is zero, the net work done in carrying a unit pole around a closed path in a magnetostatic field vanishes.

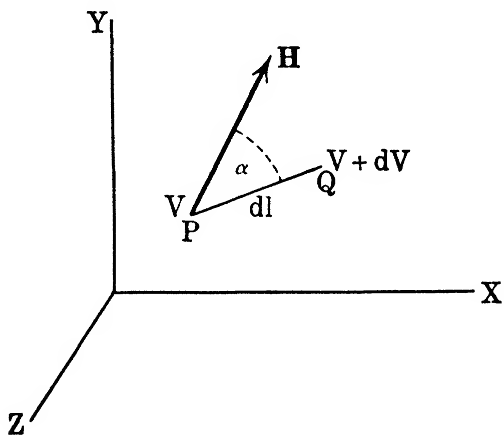


Fig. 63

If two points P and Q are very close together, as in Fig. 63, the excess of the potential of Q over that of P is

$$dV = -H \cos \alpha dl.$$

Hence the component of H in the direction of the displacement dl is

$$H \cos \alpha = -\frac{\partial V}{\partial l}. \tag{38–6}$$

As the left-hand side of this equation is greatest for α equal to zero, the potential decreases most rapidly in the direction of the magnetic intensity.

Making dl in succession equal to dx, dy and dz we find for the rectangular components of H,

$$H_x = -\frac{\partial V}{\partial x}, \qquad H_y = -\frac{\partial V}{\partial y}, \qquad = -\frac{\partial V}{\partial z}. \qquad (38\text{–}7)$$

Similarly, if V is expressed as a function of the radius vector r, polar angle θ and azimuth ϕ, the components of H in the directions of increasing r, θ and ϕ respectively are

$$H_r = -\frac{\partial V}{\partial r}, \qquad \frac{}{r\partial\theta}, \qquad \frac{}{r\sin\theta\partial\phi} \qquad (38\text{–}8)$$

Problem 38a. Three poles $-m$, $2m$, $-m$ are placed on the X axis at distances $-l$, 0, l from the origin. Find the potential at a distance r from the origin large compared to l, denoting the angle between the radius vector and the X axis by θ. Ans. $\frac{ml^2}{r^3}(1 - 3\cos^2\theta)$.

39. Gauss' Law. — This important law is deduced for a magnetostatic field in precisely the same manner as for an electrostatic field. Consider the closed surface s (Fig. 64) surrounding the point pole m. The magnetic flux dN through the surface element AB of area ds is defined as the product $H \cos \alpha ds$ of the component of H along the outward drawn normal n by the area ds of the element. The electromagnetic unit of flux is known as the *maxwell*. From (38–2) the flux through AB due to the pole m is

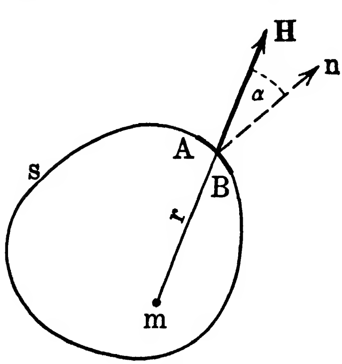


FIG. 64

$$dN = H \cos \alpha ds = m \frac{ds \cos \alpha}{}$$

But the solid angle * $d\Omega$ subtended at m by AB is

$$d\Omega = \frac{ds \cos \alpha}{r^2}.$$

Therefore

$$dN = m d\Omega,$$

and summing up over the entire surface the total outward flux is seen to be

$$N = 4\pi m. \tag{39-1}$$

On the other hand the net flux through the surface due to a pole m outside (Fig. 65) vanishes. To show this, describe a cone

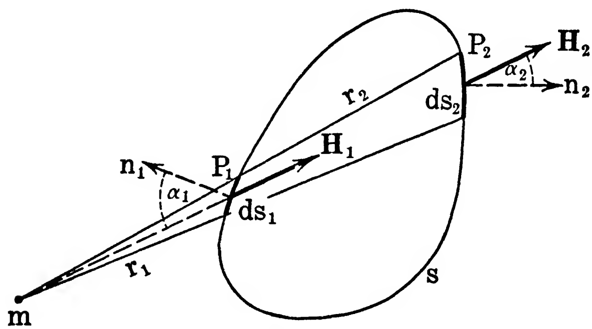


FIG. 65

of angular aperture $d\Omega$ with m as vertex. Then if ds_1 and ds_2 are the areas of the surface intercepted at P_1 and P_2,

$$d\Omega = \frac{ds_1 \cos \alpha_1}{r_1^2} = \frac{ds_2 \cos \alpha_2}{r_2^2}. \tag{39-2}$$

Taking these two elements of surface together, the flux through them is

$$\begin{aligned} dN &= H_1 \cos(\pi - \alpha_1) ds_1 + H_2 \cos \alpha_2 ds_2 \\ &= -H_1 \cos \alpha_1 ds_1 + H_2 \cos \alpha_2 ds_2 \\ &= -m \frac{ds_1 \cos \alpha_1}{r_1^2} + m \frac{ds_2 \cos \alpha_2}{r_2^2} = 0, \end{aligned}$$

* For the definition of solid angle see page 19.

on account of (39–2). As the entire surface can be divided into pairs of elements subtending the same solid angle at m, the total flux through it is zero, the inward flux through the nearer side just annulling the outward flux through the more remote side. Therefore the magnetic flux through a closed surface is due solely to the poles enclosed by the surface.

If a number of point poles $m_1, m_2, m_3, \cdots$ are inside the surface, the normal component of the resultant magnetic intensity is equal to the sum of the normal components of the magnetic intensities due to the individual poles, and therefore

$$N = 4\pi(m_1 + m_2 + m_3 + \cdots) = 4\pi\sum m. \qquad (39\text{–}3)$$

Since, however, every magnet contains equal amounts of positive and negative magnetism, a magnet entirely enclosed within a surface gives rise to no flux through it. Only when the surface cuts a magnet, so that part of the magnetic charge is inside and part outside, is the flux of H different from zero.

If magnetic charge is distributed continuously in the volume τ surrounded by the surface s with a density of ρ units of pole strength per unit volume,

$$N = 4\pi \int_\tau \rho d\tau. \qquad (39\text{–}4)$$

If we write for N the integral of the normal component of the magnetic intensity over the surface s in accord with the definition of magnetic flux, (39–3) and (39–4) take the forms

$$\int_s H \cos \alpha ds = 4\pi \sum m, \qquad (39\text{–}5)$$

and

$$\int_s H \cos \alpha ds = 4\pi \int_\tau \rho d\tau. \qquad (39\text{–}6)$$

The use of *lines of force* in describing a magnetic field is very convenient. These lines are drawn so as to have everywhere the direction of the magnetic intensity, the sense of each line being indicated by an arrow-head. A bundle of M lines of force,

where M is a large integer arbitrarily chosen, is known as a *tube of force*. Lines of force are drawn in such density that the number of tubes per unit cross-section is everywhere equal to the magnetic intensity H. By taking M large enough, the representation of a magnetic field by lines of force may be made as nearly continuous as desired, even in the case of a very weak field.

If, now, dN tubes of force pass through a small surface of area ds at an angle α with the normal to the surface, the cross-section of the tubes is $ds \cos \alpha$ and the magnetic intensity is

$$H = \frac{dN}{ds \cos \alpha}$$

or

$$dN = H \cos \alpha ds.$$

But this is just the magnetic flux through ds. Therefore *the number of tubes of force passing through a surface is equal to the magnetic flux through that surface.*

Lines of force are continuous in regions containing no magnetic poles. To prove this consider a tube bounded on the sides by lines of force and on the ends by cross-sections s_1 and s_2. Evidently there is no flux through the sides of the tube since these surfaces are loci of lines of force. Therefore the only flux is through the ends s_1 and s_2. But as there are no poles in the section of tube under consideration Gauss' law requires that the flux out of one end must equal that in through the other. Therefore as many lines of force pass out through one end of the tube as enter through the other.

By making m unity in (39–1) we see that the flux through a small surface surrounding a unit north pole is 4π. Consequently 4π tubes of force originate on every unit of positive magnetism. An equal number terminate on each unit of negative magnetism. Therefore lines of force stretch unbroken from positive magnetic charges to negative magnetic charges.

An equipotential surface is one all points of which are at the same potential. Evidently the potential decreases most rapidly in a direction at right angles to an equipotential surface. There-

fore H, since it has the direction of the most rapid decrease of potential, is everywhere perpendicular to the equipotential surfaces of a magnetic field, and as the lines of force have the direction of the magnetic intensity they intersect equipotential surfaces orthogonally.

Problem 39a. A magnet is bent so as to bring the pole pieces opposite and parallel to one another, the distance between the poles being small compared to the linear dimensions of their surfaces. If σ and $-\sigma$ are the magnetic charges per unit area of the poles, show that the magnetic intensity in the space between them is $4\pi\sigma$.

40. Magnetic Dipoles. — The molecular theory of magnetism makes the magnetic dipole or molecular magnet of fundamental importance in the study of magnetism. To find the magnetic

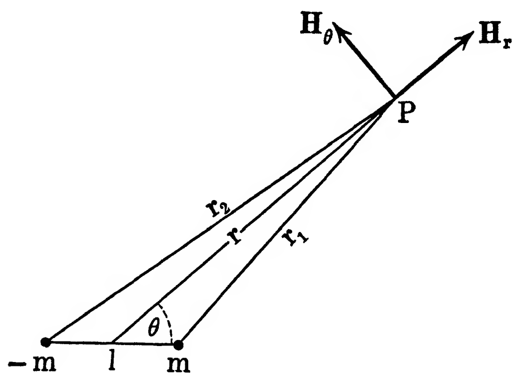


Fig. 66

potential at a point P (Fig. 66) at a distance r from the center of an elementary magnet large compared to its length l we have from (38–4)

$$V = \frac{m}{r_1} - \frac{m}{r_2} = \frac{m}{r - \dfrac{l}{2}\cos\theta} - \frac{m}{r + \dfrac{l}{2}\cos\theta} = \frac{ml\cos\theta}{r^2 - \dfrac{l^2}{4}\cos^2\theta}.$$

So

$$V = \frac{ml\cos\theta}{}.$$

as l^2 is negligible compared to r^2. The product ml is called the *magnetic moment* of the magnet and is designated by M. Evidently it is a vector, having the direction of the axis of the elementary magnet. We shall take its positive sense to be that from $-m$ to m. In terms of its magnetic moment the potential due to the elementary magnet considered above is

$$V = \frac{M \cos \theta}{r^2}. \tag{40–1}$$

The components of the field in the directions of increasing r and θ are obtained by the aid of (38–8):

$$\left.\begin{aligned} H_r &= -\frac{\partial V}{\partial r} = \frac{2M \cos \theta}{r^3}, \\ H_\theta &= -\frac{\partial V}{r\partial \theta} = \frac{M \sin \theta}{r^3}. \end{aligned}\right\} \tag{40–2}$$

The lines of magnetic force have the same configuration as the lines of electric force in the field of an electric dipole, illustrated in Fig. 25.

We have investigated the field produced by a magnetic dipole. Next we shall consider the force and torque acting on a magnetic dipole placed in an external field H.

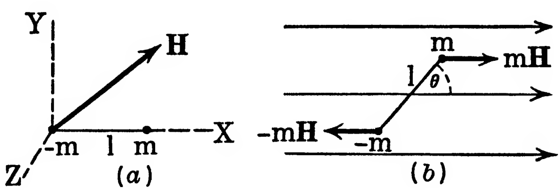


FIG. 67

To find the force take the X axis (Fig. 67*a*) parallel to the axis of the dipole. Then the X component of the force on the south pole is

$$F_{1_x} = -mH_x,$$

whereas the X component of the force on the north pole is

$$F_{2_x} = m\left(H_x + \frac{\partial H_x}{\partial x} l\right) = mH_x + M\frac{\partial H_x}{\partial x}.$$

The resultant force in the X direction is the sum F_x of F_{1_x} and F_{2_x}. In a similar manner we obtain the Y and Z components. Therefore we have

$$F_x = M\frac{\partial H_x}{\partial x}, \qquad F_y = M\frac{\partial H_y}{\partial x}, \qquad F_z = M\frac{\partial H_z}{\partial x}. \quad (40\text{–}3)$$

If the field is uniform, the components of H are constants, and the resultant force vanishes.

The torque exerted on the dipole by the field (Fig. 67*b*) is due to the couple consisting of the force mH on the north pole and $-mH$ on the south pole. As the lever arm of this couple is $l \sin \theta$, where θ is the angle between the axis of the dipole and the field, the torque is

$$L = -mHl \sin \theta = -MH \sin \theta, \qquad (40\text{–}4)$$

the minus sign indicating that the torque is in such a direction as to decrease θ. In other words the magnetic axis of the dipole tends to turn parallel to the field.

To find the energy U of the dipole due to its orientation relative to the lines of force, denote the magnetic potential at $-m$ by V_1 and that at m by V_2. Then,

$$U = -mV_1 + mV_2 = ml\left(\frac{V_2 - V_1}{l}\right) = ml\frac{\partial V}{\partial l}.$$

But

$$H \cos \theta = -\frac{\partial V}{\partial l}.$$

Hence

$$U = -mlH \cos \theta = -MH \cos \theta. \qquad (40\text{–}5)$$

Problem 40a. Two magnets of moments M_1 and M_2 are a distance r apart, the axes of the magnets lying in the line joining them. What is the force between them? Ans. $-\dfrac{6M_1M_2}{r^4}$.

Problem 40b. The second magnet of the previous problem is placed with its axis at right angles to the line joining the two. Find the torque exerted by the field of each on the other. Does the result

constitute a violation of the law of action and reaction? Why?

$$\text{Ans. } 2\frac{M_1M_2}{r^3}, \quad \frac{M_1M_2}{r^3}, \quad \text{No.}$$

Problem 40c. Deduce (40–4) from (40–5).

41. Magnetic Shells. — A *magnetic shell* is a thin sheet magnetized everywhere at right angles to its surface. The *strength* Φ of a shell is the magnetic moment per unit area of its surface. If, therefore, the shell has a magnetic charge σ per unit area on its positive side and a charge $-\sigma$ per unit area on its negative side, $\Phi = \sigma l$, where l is the thickness of the shell. The importance of the concept of the magnetic shell lies in the fact that the magnetic field produced by a current circuit is the same as that due to a shell of strength equal to the current whose periphery coincides with the circuit. This relation will be proved in Chapter VII.

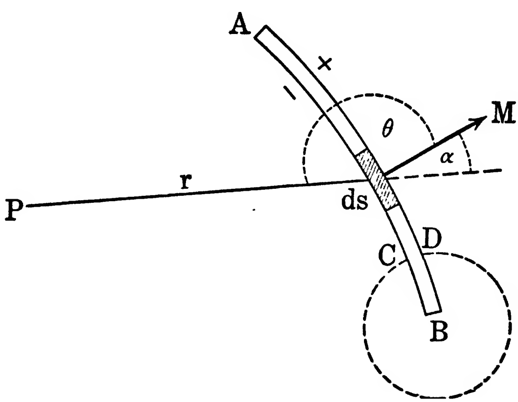


FIG. 68

Consider an element ds (Fig. 68) of the magnetic shell AB. This element is a small dipole with magnetic moment $M =$
The potential due to it at a point P is

$$\frac{\Phi ds \cos\theta}{r^2} = -\frac{\Phi ds \cos\alpha}{r^2}$$

from (40–1). Now the solid angle subtended at P by ds is

$$d\Omega = \frac{ds \cos \alpha}{r^2},$$

this angle being positive since the positive normal M to the shell makes an acute angle α with the radius vector. Therefore

$$dV = -\Phi d\Omega,$$

and the potential at P due to the entire shell is

$$V = -\int \Phi d\Omega.$$

Our chief interest is in shells of constant strength. In such a case

$$V = -\Phi\Omega, \tag{41–1}$$

where Ω is the solid angle subtended at P by the entire shell AB. Now Ω is the same for all shells having the same periphery. Consequently (41–1) tells us that all shells of the same strength and the same periphery give rise to the same potential and therefore to the same field at all outside points, exclusive, of course, of the region lying between two such shells. The shape of the shell is immaterial, the periphery and the strength being the only determining factors.

Let us calculate the work W done against the field in carrying a unit positive pole around the edge of the shell from a point C on the negative surface to an opposite point D on the positive surface of the shell. If Ω_C is the solid angle subtended at C and Ω_D that at D,

$$W = V_D - V_C = \Phi(\Omega_C - \Omega_D).$$

While Ω_C is positive, Ω_D, being the solid angle subtended at a point on the opposite side of the surface, is negative. But, as the shell is of negligible thickness, the sum of the positive magnitudes of Ω_C and Ω_D is a complete solid angle, that is, 4π. Hence

$$W = 4\pi\Phi. \tag{41–2}$$

So far we have been considering the field produced by a

magnetic shell. Let us now turn our attention to the effect of an external field H on a shell of constant strength. If θ is the angle which the positive normal to the shell makes with the lines of force, the energy of an element ds of the shell due to its position in the field is

$$dU = -\Phi ds\, H \cos\theta$$

from (40–5). Therefore the energy of the entire shell is

$$U = -\Phi \int_s H \cos\theta ds.$$

But $\int_s H\cos\theta ds$ is the magnetic flux N through the shell. Consequently

$$U = -\Phi N. \tag{41–3}$$

Let the shell suffer a displacement $d\xi$ under the action of the field. If F is the component of the force exerted on it in the direction of the displacement,

$$F d\xi = -dU = \Phi dN, \tag{41–4}$$

since the work done by F is at the expense of the energy of the shell. So if the flux through the shell is expressed as a function of the coordinates specifying its position, the force exerted on it by the field in the direction of $d\xi$ is

$$F = \Phi \frac{\partial N}{\partial \xi}. \tag{41–5}$$

In like manner the torque L corresponding to an angular displacement $d\theta$ is

$$L = \Phi \frac{\partial N}{\partial \theta}. \tag{41–6}$$

From these expressions we see that the shell tends to move in such a way as to increase the flux through it, the flux being reckoned as positive when it passes through the shell from the negative to the positive side. The shell is in stable equilibrium in the position in which the flux through it is a maximum.

Furthermore, since the flux through the shell is determined by its periphery, the energy of a shell and the force or torque acting on it depend only upon its strength and its periphery and not upon its shape.

Problem 41a. A plane shell of circular cross-section is placed at a distance r from a magnet of moment M with its center on the prolongation of the axis of the latter. The shell has a radius a and is oriented so as to be perpendicular to the radius vector r. Find the force on the shell.

$$\text{Ans. } F = -\frac{6\pi M a^2 r}{(r^2 + a^2)^{5/2}}\Phi.$$

Problem 41b. A plane shell of area A is placed in a uniform field H with its normal making an angle θ with the lines of force. Find the torque, and the period of vibration for the case where θ is small. Denote the moment of inertia of the shell by I.

$$\text{Ans. } L = -HA\Phi \sin\theta, \quad P = 2\pi\sqrt{\frac{I}{HA\Phi}}.$$

42. Density of Magnetic Charge in a Medium. — We have seen that when a magnetic medium, such as a block of iron, is placed in a magnetic field, the elementary molecular magnets are subject to a torque tending to line them up in the direction of the field. We wish to find the magnetic charge acquired by a volume element τ of dimensions Δx, Δy, Δz (Fig. 69) due to the orientation of the molecular magnets. As we can measure only the mean charge in a region containing many molecules, we shall suppose that τ is large enough to contain a great many elementary magnets. Now the charge acquired by τ is due to the passage of poles through the surfaces bounding this volume caused by the turning of the elementary magnets. Let R_x denote the *mean*

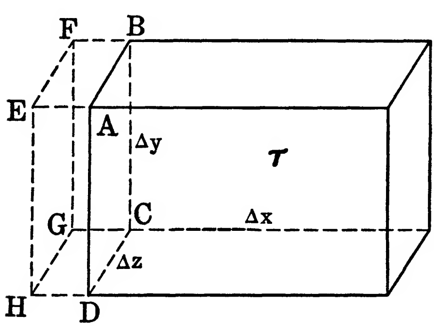


FIG. 69

displacement in the direction of the X axis of the positive poles of the molecular magnets in the neighborhood of the face $ABCD$. Then all the positive poles in a layer $ABCDEFGH$ of thickness R_x pass into τ through the face $ABCD$ due to the orientation of the elementary magnets. If, then, m is the strength of each north pole, and n the number of molecules per unit volume, the positive magnetic charge passing into τ through the surface under consideration is

$$nmR_x\Delta y\Delta z.$$

For simplicity we shall suppose that each dipole rotates about its center, the negative poles suffering displacements equal and opposite to those of the positive poles. Therefore an equal negative charge passes out of τ through $ABCD$ from a layer of thickness R_x to the right of this surface. So the net gain is

$$2nmR_x\Delta y\Delta z.$$

As R_x is measured relative to the unmagnetized state of the medium, $2R_x$ represents the X component of the mean separation of the north and south poles of the dipoles when the medium is magnetized. Consequently $2mR_x$ is the X component of the mean magnetic moment and $2nmR_x$ the X component of the resultant magnetic moment per unit volume. The magnetic moment per unit volume is known as the *intensity of magnetization* and is designated by I. It is a vector in the direction of the resultant magnetic moment of the dipoles involved. It has the same physical dimensions as H and is measured in the same units.

Now

$$I_x = 2nmR_x,$$

and the net *increase* of the magnetic charge in τ when the medium is magnetized as a result of the passage of poles through the surface $ABCD$ is

$$I_x\Delta y\Delta z. \qquad (42\text{–}1)$$

Similarly the net *decrease* due to the passage of poles through

the right-hand face parallel to $ABCD$ is

$$\left(I_x + \frac{\partial I_x}{\partial x}\Delta x\right)\Delta y\Delta z.$$

Subtracting this from the preceding expression we find

$$-\frac{\partial I_x}{\partial x}\Delta x\Delta y\Delta z$$

for the net charge acquired in so far as the two faces perpendicular to the X axis are concerned. Adding similar expressions for the two remaining pairs of faces and dividing by the volume $\Delta x\Delta y\Delta z$ of the region τ we have for the magnetic charge per unit volume due to the magnetization of the medium

$$\rho_I = -\left(\frac{\partial I_x}{\partial x} + \frac{\partial I_y}{\partial y} + \frac{\partial I_z}{\partial z}\right). \qquad (42\text{–}2)$$

This expression is identical in form with (13–1), the intensity of magnetization I taking the place of the polarization P of the electrostatic analog.

We can also express the magnetic charge acquired by a volume τ when the medium is magnetized as a surface integral. For if we replace the surface area $\Delta y\Delta z$ in (42–1) by ds and designate the angle between I and the outward drawn normal to ds by β, then

$$I_x\Delta y\Delta z = -I\cos\beta ds.$$

Consequently the net magnetic charge acquired by a volume τ when the medium is magnetized is

$$m_I = -\int_s I\cos\beta ds \qquad (42\text{–}3)$$

integrated over the surface s bounding τ. Although the charge m_I is expressed in terms of a surface integral no part of it need reside on the closed surface s. In fact it may all be located far inside the surface.

We can obtain the magnetic charge per unit area on the surface of a magnetized body by applying (42–3) to the surface of a short pill-box $ABCD$ (Fig. 70) enclosing an area s of the surface. The base AB of the pill-box lies outside the medium where I is zero, and the area of the cylindrical portion of the surface may be made so small that its contribution to the integral is negligible. Therefore (42–3) reduces to the integral over the base CD lying inside the medium. If, then, σ_I is the magnetic charge per unit area on the surface of the body,

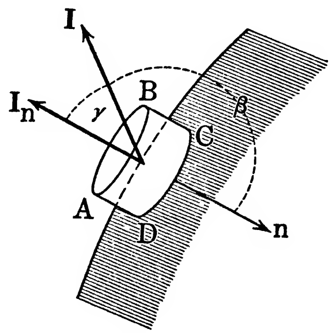


FIG. 70

$$\sigma_I s = -I \cos \beta s = I \cos \gamma s.$$

Now $I \cos \gamma$ is the component I_n of the intensity of magnetization along the outward drawn normal to the surface. Consequently

$$\sigma_I = I_n. \tag{42–4}$$

If I is perpendicular to the surface $\sigma_I = I$. For example, in the case of a uniformly magnetized bar, the magnetic charges per unit area on the ends are I and $-I$. If, however, the magnetization is not uniform the surface charge is accompanied by a volume distribution in the interior of the body given by (42–2). To calculate H in the vicinity of a magnetized body — whether outside or inside the body — we must add to the external field the field produced by the surface distribution (42–4) and the volume distribution (42–2).

Problem 42a. In actual bar magnets the intensity of magnetization is less at the ends than at the center on account of the demagnetizing effect of the poles. Suppose that the intensity of magnetization in a bar magnet of length l with its center at the origin and axis parallel to the X axis is $I = k(l^2 - x^2)$. Find the magnetic charge per unit area on the ends and per unit volume in the interior and show that the total charge is zero. Ans. $\pm \frac{3}{4}kl^2$, $2kx$.

Problem 42b. A narrow slit is cut out of a magnetic medium so that its normal makes an angle γ with I. Find the magnetic intensity in the slit due to the magnetic charges on its surface. Ans. $4\pi I \cos \gamma$.

43. Gauss' Law for a Magnetic Medium. — In magnetostatics we have no entities corresponding to free charges in electrostatics. Even the poles on the ends of a permanent steel magnet are due solely to the orientation of the elementary dipoles of which the body is composed. Therefore a magnet, as well as a paramagnetic fluid or solid which may surround it, should properly be treated as a magnetic medium.

Applying Gauss' law (39–5) to a closed surface which may surround or cut through magnets which may themselves be immersed in a paramagnetic medium

$$\int_s H \cos \alpha ds = 4\pi\Sigma m.$$

Now the only poles inside the surface s are those due to the intensity of magnetization of the media contained within this surface. The quantity Σm of this equation, then, is the m_I of (42–3), and consequently

$$\int_s (H \cos \alpha + 4\pi I \cos \beta) ds = 0. \tag{43-1}$$

If the surface s lies everywhere in the paramagnetic medium surrounding the magnets under consideration, I represents at all points of the surface the intensity of magnetization in this paramagnetic medium. If, on the other hand, the surface s cuts a magnet, I represents the intensity of magnetization in the magnet over that portion of s lying inside the magnet.

The vector sum $\mathbf{H} + 4\pi\mathbf{I}$ is known as the *magnetic induction* $\mathbf{B}$, that is,

$$\mathbf{B} \equiv \mathbf{H} + 4\pi\mathbf{I}. \tag{43-2}$$

Evidently B is measured in the same units as H and I.

As B is the vector resultant of H and $4\pi I$ the component of B normal to a surface element ds is equal to the sum of the normal

components of H and $4\pi I$. So if the angle which B makes with the normal is denoted by γ,

$$B \cos \gamma = H \cos \alpha + 4\pi I \cos \beta,$$

and Gauss' law becomes

$$\int_s B \cos \gamma ds = 0. \qquad (43\text{–}3)$$

If we define the *flux of induction* through a surface ds as the product of the component of B along the outward drawn normal by the area of the surface, Gauss' law (43–3) states that the total outward flux of induction through *any* closed surface is equal to zero.

Lines of induction may be drawn in a magnetic field so as to be everywhere parallel to B and in such density that the number of *tubes of induction* — a tube being a bundle of a specified number of lines — per unit cross-section is equal to the magnitude of B. The flux of induction through any surface is equal to the number of tubes of induction passing through the surface, the proof of this statement being identical with that of the corresponding relation between flux of force and tubes of force considered in article 39. If we consider a tube bounded by lines of induction and terminated by cross-sections s_1 and s_2, Gauss' law requires that the flux out of one end of the tube must equal that in through the other end. Therefore as many lines of induction must pass out of the one end as enter the other. It follows that lines of induction are everywhere continuous. They form closed curves, never terminating. In particular, lines of induction are continuous when we pass from one magnetic medium to another or from the interior of a magnet into the space surrounding it. In empty space B and H are the same and there is no distinction between lines of force and lines of induction.

In the case of isotropic paramagnetic media in such fields as are ordinarily available the state of saturation is far from being realized and the intensity of magnetization I is in the same direction as and proportional to the field H. Hence we may write

$$I = \epsilon H, \qquad (43\text{–}4)$$

the constant ϵ being known as the *magnetic susceptibility*. Therefore (43–2) gives

$$B = H + 4\pi I = \mu H, \tag{43-5}$$

where the *permeability* μ is defined by

$$\mu \equiv 1 + 4\pi\epsilon. \tag{43-6}$$

While μ is constant for paramagnetic substances in available fields, a general relation of the form $B = \mu H$ holds only as a rough approximation in the case of ferromagnetic media where, on account of hysteresis, B is not even a one-valued function of H.

We are now ready to compute the field produced by a pole m (Fig. 71) on the end of a long needle immersed in a paramagnetic medium, such as oxygen. We shall suppose that the pole of opposite sign is too far away to produce any appreciable effect in the neighborhood of m. Describe a sphere of radius r about m as center and apply Gauss' law (43–3) to its surface. Let B be the magnetic induction in the surrounding medium and B' that in the needle NS. We have to evaluate the integral involved in Gauss' law in two parts, first writing down the integral over the portion of the spherical surface which lies in the paramagnetic medium and then adding on that over the portion of the surface which lies inside the needle. If we denote the cross-section of the needle by A we have then

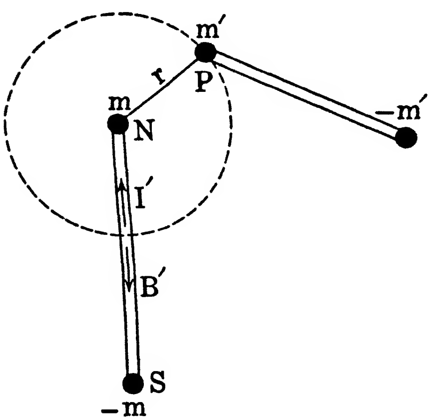


Fig. 71

$$(4\pi r^2 - A)B + AB' = 0.$$

The intensity of magnetization I' inside the needle is directed toward the center of the sphere. So, taking account of signs,

$$B' = H - 4\pi I',$$

and

$$(4\pi r^2 - A)B + AH - 4\pi AI' = 0.$$

Now AI' is the strength m of the pole on the end of the needle. So

$$4\pi r^2 B - A(B - H) = 4\pi m.$$

By making the needle thin enough the term in A can be made negligible. Hence

$$4\pi r^2 B = 4\pi m. \tag{43-7}$$

As the left-hand side of this equation represents the number of tubes of induction passing out through the portion of the surface of the sphere which lies outside the needle, we note that 4π tubes of induction must pass out through a unit positive point pole into the medium surrounding it. Since, however, tubes of induction are continuous, the same number of tubes must pass into the unit positive pole under consideration through the interior of the needle.

Replacing B by μH in the last equation, where μ is the permeability of the surrounding medium,

$$H = \frac{m}{\mu r^2}. \tag{43-8}$$

This expression represents the mean force per unit pole on a small positive pole placed in the interstices between the atoms of the medium at a distance r from m.

It follows from (43-8) that the expressions (38-3) and (38-4) for the potential become

$$V = \frac{m}{\mu R}, \tag{43-9}$$

and

$$V = \sum \frac{m}{\mu r}, \tag{43-10}$$

when the magnetic charges are immersed in a paramagnetic medium.

Just as in the electrostatic analog treated in article 16, a

paramagnetic medium in a magnetic field is subject to stresses due to the forces acting on the elementary magnetic dipoles. In order to apply the formulas of article 16 to the magnetic case it is only necessary to replace D by B, E by H and κ by μ. Thus the pressure in a paramagnetic medium is

$$p = \frac{\mu - 1}{8\pi} H^2 = \frac{1}{8\pi}\{BH - H^2\} \tag{43-11}$$

from (16–3), and from (16–7) the tension stress on a surface of the medium lying at right angles to the lines of force is seen to be

$$S = \frac{\mu - 1}{\mu}\frac{H_0{}^2}{8\pi} \tag{43-12}$$

in terms of the field H_0 outside the permeable medium. The latter should give rise to an increase in length of a solid paramagnetic rod placed parallel to the lines of force, but the permeabilities of paramagnetic media are so nearly unity that this effect has not been detected with certainty. On the other hand ferromagnetic materials exhibit a considerable change in length when placed in a magnetic field, a phenomenon known as *magnetostriction*. The effect is, however, often a shortening instead of a lengthening, and is undoubtedly caused more by mechanical stresses produced by the orientation of the elementary magnetic dipoles than by the magnetic stresses under discussion here.

To return to Fig. 71, suppose that a second magnetic needle is placed with its north pole m' at P. For the moment let us consider this pole to be a small sphere with the magnetic charge distributed over its surface in such a manner as to annul the field H in its interior. The sphere constitutes the magnetic analog of an electrical conductor, and from (16–11) we have for the force exerted by m on m'

$$F = \frac{mm'}{\mu r^2}. \tag{43-13}$$

If, now, we decrease the dimensions of m' until it becomes effectively a point pole, this expression for the force, since it does

not involve the dimensions of the sphere, remains valid. Consequently (43–13) represents the force between two point poles a distance r apart in a paramagnetic medium.

Problem 43a. Show that B inside a magnetic medium is the force that would be experienced by a unit north pole placed in a narrow slit cut so that its faces are perpendicular to I, and that H is the force at the center of a needle like cavity cut with its axis parallel to I. (Kelvin's definitions of B and H.)

Problem 43b. A long bar magnet is uniformly magnetized with intensity of magnetization I. Find B and H (1) just outside the north pole, (2) just inside the north pole, (3) at the center of the magnet. Note that B is twice as great at the center of the magnet as at its ends. Ans. (1) $2\pi I$, $2\pi I$; (2) $2\pi I$, $-2\pi I$; (3) $4\pi I$, 0.

Problem 43c. Near the middle of the magnet of the previous problem is B the same at a point outside the bar as inside? Ans. No. (0 and $4\pi I$).

44. Uniformly Magnetized Sphere. — The field produced by a uniformly magnetized sphere can be obtained easily by elementary methods. Consider concentric spheres of equal radius a one of which has a uniform volume density ρ of positive magnetic charge and the other an equal volume density of negative charge. If the first is displaced a small distance l (Fig. 72) relative to the second every element of positive charge is displaced a distance l from the negative charge with which it originally coincided, forming therewith an elementary magnetic dipole. By this displacement we have arrived at a uniformly magnetized sphere of intensity of magnetization $I = \rho l$.

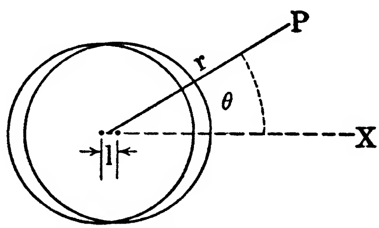


Fig. 72

Now the field at outside points due to a spherical distribution of magnetic charge is the same as if the entire charge were concentrated at the center of the sphere. The proof of this statement follows from Gauss' law in exactly the same manner as in the case of the electrostatic analog treated in article 9. Con-

sequently the magnetic field at P is that produced by a dipole of moment

$$M = \frac{4}{3}\pi a^3 \rho l = \frac{4}{3}\pi a^3 I \tag{44-1}$$

located at the origin. The potential outside the sphere is given by (40–1) and the radial and transverse components of the magnetic intensity by (40–2).

To find the field inside the sphere we must remember that the magnetic intensity at a point distant r from the center of a sphere which has a uniform distribution of magnetic charge is that produced by the charge inside a sphere of this radius. Following the same method as was employed for the electrostatic case in article 9, we have from (9–2)

$$H = \frac{m}{a^3} r,$$

if we replace Q by the magnetic charge $m = (4/3)\pi a^3 \rho$. As H is directed along the radius vector, the vector fields $\mathbf{H}_1$ and $\mathbf{H}_2$ due to the positive and negative spheres of Fig. 72 are proportional to the vector distances $\mathbf{r}_1$ and $\mathbf{r}_2$ (Fig. 73) from their respective centers, the first being directed along $\mathbf{r}_1$ and the second, since the charge is negative, in the direction opposite to $\mathbf{r}_2$. But the vector sum of $\mathbf{r}_1$ and $-\mathbf{r}_2$ is $-\mathbf{l}$, the separation of the two spheres. Therefore the resultant field is in the negative X direction and has the magnitude

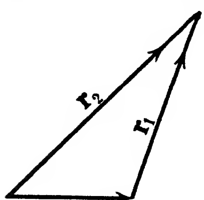


FIG. 73

$$H_x = -\frac{m}{a^3} l = -\frac{4}{3}\pi \rho l = -\frac{4}{3}\pi I. \tag{44-2}$$

This field is uniform throughout the interior of the sphere. Being in the opposite direction to I it tends to weaken the magnetization of the sphere. It is known as a *demagnetizing field*. The presence of the demagnetizing field means that the intensity of magnetization of an isolated bar magnet of ferromagnetic material is always less than the retentiveness Oc

(Fig. 60) which would represent the magnetization if no poles existed on its ends.

The magnetic induction inside the sphere is

$$B_x = H_x + 4\pi I = \frac{8}{3}\pi I, \tag{44-3}$$

in the positive X direction.

So far we have been considering a permanent magnet in the form of a uniformly magnetized sphere, such as a spherical steel magnet. Let us now consider a paramagnetic sphere of permeability μ placed in a uniform external field H_0 parallel to the X axis. As the field H_0 is uniform, the elementary dipoles in the interior of the sphere will be in equilibrium if the sphere becomes uniformly magnetized by induction so as to produce a uniform field H of its own of precisely the same character as that of the permanent spherical magnet just considered. The total field is the resultant of H_0 and H. The components of the latter at points outside the sphere are given by (40–2). Therefore the resultant field is

$$H_r = H_0 \cos\theta + \frac{8\pi a^3 I}{3r^3}\cos\theta \tag{44-4}$$

in the direction of the radius vector, and

$$H_\theta = -H_0 \sin\theta + \frac{4\pi a^3 I}{3r^3}\sin\theta \tag{44-5}$$

at right angles thereto in the direction of increasing θ.

Inside the sphere the field is in the X direction and equal to the sum of H_0 and the internal field (44–2). It is

$$H_x = H_0 - \frac{4\pi I}{3}. \tag{44-6}$$

If we apply Gauss' law in the manner of article 14 we find that when we pass from one magnetic medium to another the components of B normal to the surface of separation and the components of H parallel to the surface are the same in the two media. The proof is identical with that given for the electro-

static analog in article 14, B taking the place of D and H that of E.

Therefore we have from (44–4) and (44–6)

$$\left(H_0 + \frac{8\pi I}{3}\right)\cos\theta = \mu\left(H_0 - \frac{4\pi I}{3}\right)\cos\theta$$

for the normal components of B, or

$$I = \frac{3}{4\pi}\frac{\mu - 1}{\mu + 2}H_0. \qquad (44\text{–}7)$$

Equating the tangential components of H gives no new information since an identity is obtained.

Equation (44–7), then, specifies the intensity of magnetization produced in the paramagnetic sphere by the impressed field H_0. The student should compare this result with that obtained in problem 28*b* for a dielectric sphere placed in a uniform electric field.

Substituting (44–7) in (44–6) the resultant field in the interior of the sphere is seen to be

$$H_x = \left(1 - \frac{\mu - 1}{\mu + 2}\right)H_0, \qquad (44\text{–}8)$$

which is less than H_0 if μ is greater than unity, due to the fact that the induced poles on the surface of the sphere give rise to a field in the interior opposed to H_0. This field is the demagnetizing field. For the case under consideration it has the magnitude

$$H_x' = -\frac{\mu - 1}{\mu + 2}H_0 = -\frac{4\pi}{3}I. \qquad (44\text{–}9)$$

The ratio of the magnitude of the demagnetizing field H_x' to that of the intensity of magnetization I is sometimes called the *demagnetizing factor*. Our analysis shows that its value for a sphere is $4\pi/3$.

The magnetic induction inside the sphere is

$$B_x = \mu H_x = \frac{3\mu}{\mu + 2}H_0 = \frac{3}{1 + \dfrac{2}{\mu}}H_0, \qquad (44\text{–}10)$$

which is greater or less than H_0 according as μ is greater or less than unity. Therefore lines of induction are crowded together in a paramagnetic sphere as illustrated in Fig. 52. In a diamagnetic sphere, on the contrary, $\mu < 1$ and the lines of induction are spread apart.

Problem 44a. As a rough approximation the earth may be considered to be a uniformly magnetized sphere. If the horizontal field strength at the equator is 0.5 gauss, what is the intensity of magnetization? Ans. 0.12 e.m.u.

Problem 44b. A long paramagnetic cylinder of circular cross-section is placed in a uniform magnetic field H_0 with its axis at right angles to the lines of force. By the method of this article show that the magnetic induction in the cylinder is

$$B = \frac{2}{1 + \frac{1}{\mu}} H_0,$$

and that the demagnetizing factor is 2π.

Problem 44c. Solve the problem of the paramagnetic sphere, treated in this article, by the use of zonal harmonics.

45. Magnetic Shielding. — A sensitive magnetic instrument can be shielded very effectively from outside fields by placing it inside a cylindrical shell made of soft iron of high permeability. We shall investigate the field in the interior of a paramagnetic cylindrical shell placed in a uniform magnetic field at right angles to the axis of the cylinder.

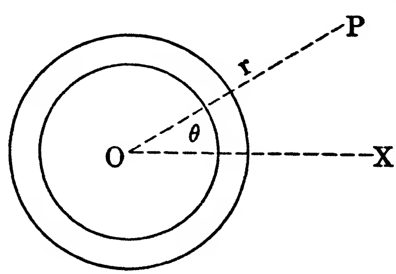


FIG. 74

As the same inverse square law of attraction and repulsion holds for magnetic as for electric charges, the magnetic potential in a region containing no magnetic charge must satisfy Laplace's equation (26–2) for the same reason that this equation is satisfied by the electric potential in a region containing no electric charge. Consider, now, an infinitely long cylindrical shell (Fig. 74) of inner radius a and outer radius b placed in a uniform magnetic

field H_0 parallel to the X axis. As in the case of a homogeneous isotropic dielectric (art. 15) there can be no volume distribution of charge inside a magnetic medium in which I is proportional to H. Consequently we need three solutions of Laplace's equation; a solution V_1 to represent the potential outside the shell, a second solution V_2 to represent the potential in the material of the shell, and a third V_3 to represent the potential in the cavity. Evidently cylindrical harmonics are indicated, the potential being a function of r and θ alone.

The boundary conditions to be satisfied are the following. (*a*) For r very large V_1 must become $-H_0x = -H_0r\cos\theta$. (*b*) At each surface of the shell the radial component of the magnetic induction and the transverse component of the magnetic intensity must be continuous as we pass from the outside to the inside of the shell. The latter of these two conditions is evidently satisfied if the potential itself is continuous.

As the potential at infinity must be $-H_0r\cos\theta$ it is clear that we are limited to harmonics involving only the first power of $\cos\theta$. Referring to Table II, page 90, we see that there are only two cylindrical harmonics satisfying this condition and that the three potentials must be of the form

$$\left.\begin{aligned} V_1 &= -H_0r\cos\theta + \frac{B_1\cos\theta}{r}, \\ V_2 &= A_2r\cos\theta + \frac{B_2\cos\theta}{r}, \\ V_3 &= A_3r\cos\theta, \end{aligned}\right\} \quad (45\text{-}1)$$

where we have omitted the term in $1/r$ from V_3 since its presence would make the potential infinite at the origin.

In order that $V_1 = V_2$ for $r = b$ it is necessary that

$$-H_0b^2 + B_1 = A_2b^2 + B_2, \qquad (45\text{-}2)$$

and in order that $V_2 = V_3$ for $r = a$,

$$A_2a^2 + B_2 = A_3a^2. \qquad (45\text{-}3)$$

Continuity of the radial component of the magnetic induction

$$\mu H_r = -\mu \frac{\partial V}{\partial r}$$

at the outer surface requires that

$$H_0 b^2 + B_1 = \mu(-A_2 b^2 + B_2), \tag{45-4}$$

and at the inner surface that

$$\mu(-A_2 a^2 + B_2) = -A_3 a^2. \tag{45-5}$$

In equations (45–2) to (45–5) we have four relations from which we can eliminate B_1, A_2, B_2 so as to find A_3 in terms of H_0. Putting the value of A_3 so found into the expression (45–1) for V_3,

$$V_3 = -\frac{4\mu}{(\mu+1)^2 - \frac{a^2}{b^2}(\mu-1)^2} H_0 r \cos\theta.$$

The ratio of the field H_0 existing before the introduction of the shell to the uniform field

$$H_3 = -\frac{\partial V_3}{\partial x} = \frac{4\mu}{(\mu+1)^2 - \frac{a^2}{b^2}(\mu-1)^2} H_0$$

in the cavity is called the *shielding ratio*. It is

$$g = \frac{H_0}{H_3} = \frac{1}{4\mu}\left\{(\mu+1)^2 - \frac{a^2}{b^2}(\mu-1)^2\right\}$$

$$= 1 + \frac{1}{4}\frac{(\mu-1)^2}{\mu}\left(1 - \frac{a^2}{b^2}\right). \tag{45-6}$$

If the permeability of the shell is large compared to unity, as in the case of soft iron where it may be several hundred, the shielding ratio is given very closely by

$$g = \frac{1}{4}\mu\left(1 - \frac{a^2}{b^2}\right). \tag{45-7}$$

If m is the mass of the cylindrical shell per unit length and D its density, $m = \pi D(b^2 - a^2)$. Consequently we may write

(45–7) in the form

$$g = \frac{m\mu}{4\pi D b^2}. \tag{45–8}$$

This expression shows that, for a given mass of soft iron, the shielding ratio is greater the smaller the outside radius of the cylindrical shell.

The effectiveness of magnetic shielding is somewhat increased by using several coaxial cylindrical shells with air gaps between. Wills has analyzed the case of three such shells and finds that for a given mass of iron the most effective arrangement is that in which the radii of the successive surfaces form a geometrical progression.

Problem 45a. A cylindrical shell made of iron of permeability 300 has radii of 4 cm and 5 cm. What is the shielding ratio? Ans. 27.

Problem 45b. Find the radial and transverse components of the intensity of magnetization inside the cylindrical shell of the last article for μ large. Sketch roughly the lines of induction.

$$\text{Ans. } I_r = \frac{H_0}{2\pi}\,\frac{1 - \dfrac{a^2}{r^2}}{1 - \dfrac{a^2}{b^2}}\cos\theta,$$

$$I_\theta = -\frac{H_0}{2\pi}\,\frac{1 + \dfrac{a^2}{r^2}}{1 - \dfrac{a^2}{b^2}}\sin\theta.$$

46. Energy in the Magnetic Field. — Since any system of magnets may be built up of elementary magnetic dipoles (art. 40), a certain amount of work being performed in the process, it is evident that such a system has a potential energy associated with it. That is, during the establishment of the given configuration an amount of energy is stored which may be obtained again on scattering the magnetic elements to infinity. As we are dealing with magnetostatics, it is only the final state of the system that determines the potential energy. We may build up the system in any manner convenient for calculation. For

simplicity let us consider the case of several long thin fixed magnets immersed in a medium of permeability μ, the pole strengths being m_1, $-m_1$, m_2, $-m_2$, and so on. These poles are built up step by step by bringing magnetic charge from infinity in such a way that at any instant all poles are in magnitude the same fraction of their final values. That is, during the building process the pole strengths are $m_1\alpha$, $-m_1\alpha$, $m_2\alpha$, $-m_2\alpha$, and so on, where α varies from 0 to 1. Furthermore, under the above conditions the potential at every point varies in the same way as the pole strengths. Thus, if V is the final value of the potential at any one of the poles m, the potential is $V\alpha$ when the pole strength is $m\alpha$. This follows at once from (43–9). A single step in the magnetizing process consists of increasing every pole $m\alpha$ by an amount $md\alpha$, since by so doing all poles are kept the same fraction $(\alpha + d\alpha)$ of their final value. By definition the potential at a pole is the work done against the magnetic forces in bringing a unit positive pole from infinity. Hence the work done in the step described above is given by

$$dU = \Sigma V\alpha m d\alpha = \alpha d\alpha \Sigma m V,$$

the summation being taken over all poles, positive and negative, with appropriate sign. Thus the total energy stored during the establishment of the final configuration is

$$U = \left\{ \int_0^1 \alpha d\alpha \right\} \Sigma m V = \tfrac{1}{2} \Sigma m V. \qquad (46\text{–}1)$$

As this energy may be obtained again by scattering the elements of the fixed magnets to infinity it is *potential energy.* The location of the energy is to some extent indeterminate, as is also the case with electrostatic energy (art. 21). However, when there is a permeable medium present some of the energy at least is distributed through it, on account of the induced magnetization. It is convenient, therefore, to consider all the energy as distributed through the medium, even when the latter consists only of free space. This may be done by observing that the tubes of induction, which are continuous, pass through

the given point poles to the number of 4π per unit pole, as shown in article 43. Consider the portion of a tube of induction running from $1/4\pi$ of a unit pole on the positive end of some magnet to $-1/4\pi$ of a unit pole on some negative end. The energy which is associated with it is evidently

$$\frac{1}{2}\left(\frac{1}{4\pi}\right)V_a+\frac{1}{2}\left(-\frac{1}{4\pi}\right)V_b$$
$$=\frac{1}{8\pi}(V_a-V_b)=\frac{1}{8\pi}\int_a^b Hdl, \quad (46\text{–}2)$$

where V_a and V_b are the potentials at the beginning and end of the section of tube considered. This energy we may distribute along the tube as we choose. The simplest way is at the rate of $H/8\pi$ ergs per centimeter at every point along the given section of tube. Since there are B tubes per square centimeter of cross-section (art. 43) energy is distributed through the field in the amount of $BH/8\pi$ ergs per cubic centimeter. Replacing B by μH we have for a paramagnetic medium

$$U=\int_\tau \frac{\mu H^2}{8\pi}d\tau, \quad (46\text{–}3)$$

where the volume integral is taken through all space, since the volume occupied by the long thin magnets is negligible.

The distribution of energy specified in (46–3) is an arbitrary one, since we may add to $\mu H^2/8\pi$ any function whose integral through space is zero without changing the total energy. It is however the most reasonable distribution, as it assigns to the induced magnetization exactly the correct energy per unit volume. For

$$\frac{\mu H^2}{8\pi}=\frac{H^2}{8\pi}+\frac{1}{2}\epsilon H^2, \quad (46\text{–}4)$$

the term involving the susceptibility ϵ representing the energy in question.

Now, on the hypothesis that magnetization consists in rotating

elementary magnets, the work done may be calculated at once. Thus if a magnet (Fig. 75) of length l and pole-strength m rotates through an angle $d\theta$ under the influence of a field H, the work done is

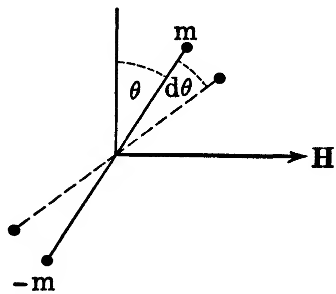


FIG. 75

$$Hml \cos \theta d\theta = HdM_H,$$

where M_H is the component of the magnetic moment parallel to the field. The work done per unit volume is then HdI, so that the total work of magnetization is given by

$$\int_0^I HdI = \epsilon \int_0^H HdH = \tfrac{1}{2}\epsilon H^2,$$

which agrees with (46–4).

If some part of the magnetic system, say one of the fixed magnets, is allowed to move a distance $d\xi$ the force is, as in (41–4),

$$F = -\frac{\partial U}{\partial \xi}. \qquad (46\text{–}5)$$

Similarly for a rotation the torque is

$$L = -\frac{\partial U}{\partial \theta}. \qquad (46\text{–}6)$$

Problem 46a. A fixed magnet of large cross-section and uniform magnetization I_0 is bent around until its pole faces are parallel, with a separation d. These pole faces are planes perpendicular to the direction of magnetization. Between the faces is a plate of soft iron of thickness t and permeability μ. It is withdrawn until only a length x remains between the poles. Find the force per cm of width tending to draw the plate back to its original position.

Ans. $2\pi I_0^2 t \dfrac{\mu - 1}{\mu}$.

47. Magnetic Instruments and Measurements. — One of the most useful magnetic instruments is the *magnetometer* (Fig. 76). It appears in a variety of special forms, but consists essentially of a small magnet suspended by a long torsionless fibre in a non-magnetic case attached to the base of which are two horizontal arms with sliding cradles. In these may be placed small auxiliary magnets which deflect the suspended magnet from its normal equilibrium position in the magnetic meridian of the earth's field. The suspended system usually carries a small mirror, so that the deflection may be observed by means of an external telescope and scale, or some such device.

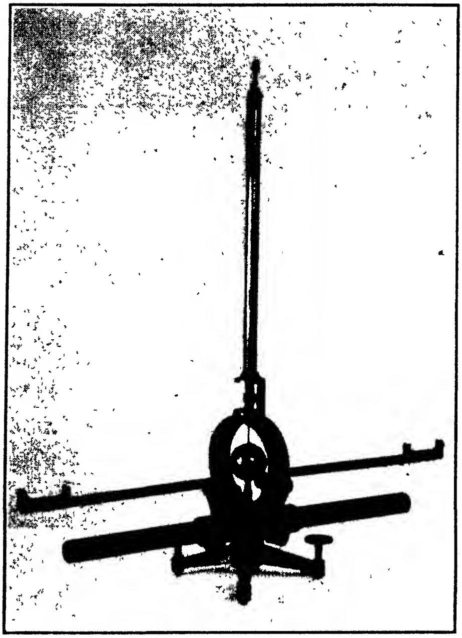


FIG. 76

To measure the horizontal component H_e of the earth's field, we first determine the magnetic meridian and set the arms in a direction perpendicular to it. A small magnet of unknown moment M (Fig. 77) is placed in one of the cradles with its axis perpendicular to the meridian and its center a distance d from the center of the suspended magnet, whose moment is M'. The movable system is deflected through an angle θ. Let H_m be the field of M at the center of M'. As M' is very short the variation of field over the region occupied by M' may be neglected. Then θ is determined by the equation

$$H_m M' \cos\theta - H_e M' \sin\theta = 0, \qquad (47\text{–}1)$$

obtained by equating the resultant torque on M' to zero. If now M is so small that it may be regarded as a magnetic dipole at distance d, then, by (40–2),

$$H_m = \frac{2M}{d^3},$$

and (47–1) reduces to

$$\frac{M}{H_e} = \frac{d^3}{2} \tan \theta. \tag{47–2}$$

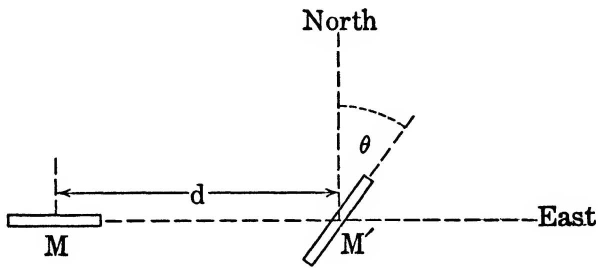


FIG. 77

In order to eliminate the unknown M we must find another relation between M and H_e. This is conveniently done by suspending M in the earth's field and allowing it to oscillate in a horizontal plane about an axis perpendicular to its length passing through its center of mass. A torsionless fibre is again used, and the oscillations are so small that the motion is effectively simple harmonic. The equation of motion is then

$$I\frac{d^2\phi}{dt^2} = -\,MH_e\phi,$$

where I is the moment of inertia about the given axis, and ϕ is the angular deflection. The period is given by

$$P_0 = 2\pi\sqrt{\frac{I}{MH_e}}. \tag{47–3}$$

Eliminating M between this equation and (47–2) gives

$$H_e = \frac{2\pi}{P_0}\sqrt{\frac{2I}{d^3 \tan \theta}}. \tag{47–4}$$

Evidently we may eliminate H_e and obtain

$$M = \frac{2\pi}{P_0}\sqrt{\frac{Id^3 \tan\theta}{2}}. \qquad (47\text{–}5)$$

Unfortunately in practice the value of θ is usually too small for accurate measurement if d is made so large that the length of M is negligible. In fact the ratio of these quantities is rarely as great as ten. Hence we require a more accurate expression for H_m. Referring again to article 40 we see that if we do not neglect $l^2/4$ in comparison with d^2,

$$H_m = -\left[\frac{\partial}{\partial r}\left(\frac{ml}{r^2 - \frac{l^2}{4}}\right)\right]_{r=d} = \frac{2mld}{\left(d^2 - \frac{l^2}{4}\right)^2} \doteq \frac{2M}{d^3}\left(1 + \frac{l^2}{2d^2}\right),$$

where l is the distance between the poles of M, a quantity in general slightly less than the geometrical length of the magnet. Equation (47–2) is now replaced by

$$\frac{M}{H_e}\left(1 + \frac{l^2}{2d^2}\right) = \frac{d^3}{2}\tan\theta. \qquad (47\text{–}6)$$

As l cannot be measured exactly, it is eliminated by making two sets of observations. Thus let θ_1 correspond to d_1, and θ_2 to d_2. Then

$$\frac{M}{H_e}\left(1 + \frac{l^2}{2d_1^2}\right) = \frac{d_1^3}{2}\tan\theta_1,$$

$$\frac{M}{H_e}\left(1 + \frac{l^2}{2d_2^2}\right) = \frac{d_2^3}{2}\tan\theta_2,$$

and, on eliminating l^2,

$$\frac{M}{H_e} = \frac{d_1^5\tan\theta_1 - d_2^5\tan\theta_2}{2(d_1^2 - d_2^2)}. \qquad (47\text{–}7)$$

This last equation is used with (47–3) to obtain H_e.

In order to avoid errors due to lack of magnetic and mechanical symmetry, the auxiliary magnet is turned end for end, and then transferred from the west arm to the east arm and the process repeated. The mean of the four observations gives the value of θ to be used. Also, care must be taken that the magnetometer fibre is really torsionless, or that the effect of any

existing torsion is removed by turning the upper end of the fibre through the same angle as the lower for each deflection. In the *Kew* method of making the observations this result is obtained automatically by rotating the entire instrument.

Similarly, in the oscillation observations various refinements and corrections are necessary. For example, the period can be determined accurately only by observing the time for a large number of vibrations. Because of the mechanical damping the initial amplitude must then be too large strictly to satisfy the condition for simple harmonic motion. If P is the average period determined for 100 vibrations, say, and α_1, α_2 are the amplitudes in radians at the beginning and the end of the interval over which P is determined, then approximately

$$P_0 = P\left(1 - \frac{\alpha_1\alpha_2}{16}\right), \tag{47–8}$$

provided α_1 and α_2 are not small compared to the difference between them. Corrections for temperature effects, residual torsion in the fibre, and variation of M due to H_e are of somewhat less importance. For a description of these the reader is referred to a more elaborate discussion of magnetometer measurements. (See, for example, Glazebrook: *Dictionary of Applied Physics*, Vol. II, p. 532 ff.)

It is interesting to note that Gauss used the magnetometer to demonstrate the validity of the inverse square law (37–1). Considering the arrangement shown in Fig. 77 we obtain

$$\frac{M}{H_e} = \frac{d^3}{2}\tan\theta, \tag{47–9}$$

as has been shown. Suppose now M is moved parallel to itself in the horizontal plane until its center lies in the meridian at a distance d again from the center of M'. Then, by (40–2),

$$H_m = \frac{M}{d^3},$$

and

$$\frac{M}{H_e} = d^3 \tan\theta'. \tag{47–10}$$

Thus,

$$\frac{\tan\theta}{\tan\theta'} = 2. \tag{47-11}$$

On the other hand it may be shown that if the force between magnetic poles varies as $1/r^p$ we must have

$$\frac{\tan\theta}{\tan\theta'} = p.$$

As the ratio of the tangents is found experimentally to be 2, the inverse square law is confirmed.

Another instrument of use in the measurement of terrestrial magnetism is the *dip-circle* (Fig. 78). It consists of a light magnetic needle arranged to rotate about a horizontal axis, and a circular scale over which the ends of the needle move. The axis of rotation being set perpendicular to the magnetic meridian, the needle assumes a position tangent to the lines of force of the earth's field. The angle which the needle makes with the horizontal, that is, the *angle of dip*, is determined from the position of the ends of the needle on the scale. As in the case of the magnetometer, various precautions are necessary to avoid errors due to lack of mechanical and magnetic symmetry in the instrument. A detailed discussion of the use of the dip-circle is given in Glazebrook, in the article on *Terrestrial Magnetism* to which reference is made above.

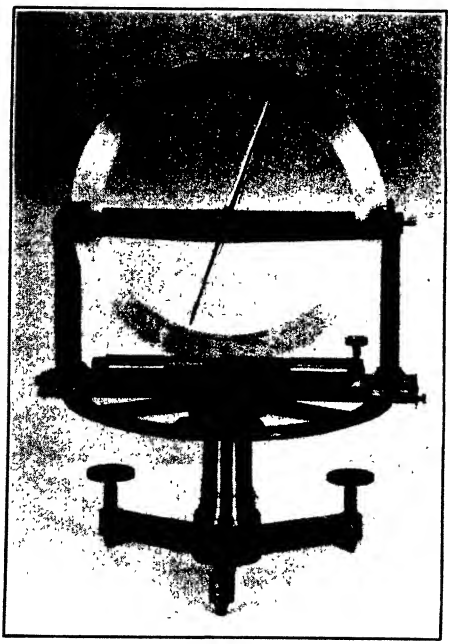


FIG. 78

CHAPTER V

STEADY CURRENTS

48. Current and Electromotive Force. — A free charge placed in an electric field experiences a force in the direction of the field, and moves in accordance with the usual laws of mechanics. Let us place a conductor in the field. As the conductor contains a great number of free charges and the field, initially at least, penetrates into its interior, a flow of electricity takes place. If the conductor is isolated the charge soon distributes itself in such a way as to make the field within the conductor everywhere zero, and motion of charge then ceases. This static state is considered in earlier chapters, and we are not concerned with it here. If, on the other hand, the conductor is connected to reservoirs which supply or absorb charge as required to maintain the original electric field, charge passes continuously through the conductor. As there can be no continued accumulation of charge at any point along the conductor, the amount of charge passing per second is the same for every cross-section which cuts all the lines of flow. Moreover, although each free charge is continuously accelerated by the field, it progresses through the conductor at a constant speed on the average due to the mechanical retardation it experiences from frequent collisions with the atoms of which the conductor is composed. The total charge passing any cross-section per second is the *electric current* in the conductor. We will denote it by i. Suppose the conductor is a wire of cross-section A. If there are n free charges per unit volume, each with charge e and average speed of progression v, then

$$i = Anev. \tag{48–1}$$

The current j per unit area of cross-section, called the *current density*, is given by

$$j = \frac{i}{A} = nev. \tag{48–2}$$

The positive direction of the current is that of the electric field, the direction in which positive charges tend to move. As the free charges in a conductor are electrons, with $e = -4.77(10)^{-10}$ e.s.u., the actual direction of flow is opposite to the conventional direction of the current.

Evidently the magnitude of the current between two points depends on the nature of the conductor and on the forces which urge the free charges through it. The effect of all such forces is most easily calculated in terms of the *work performed per unit positive charge.* This quantity is called the *electromotive force*, often abbreviated e.m.f., and is denoted by $\mathcal{E}$. Thus, suppose we carry a unit positive charge along some path from P to Q (Fig. 79) in a field of force. If E is the magnitude of the total or resultant force per unit positive charge, and α is the angle which the force makes with an element dl of the path, we see that

$$\mathcal{E} = \int_P^Q E \cos \alpha dl, \tag{48-3}$$

for this is the work done by the field on the unit charge.

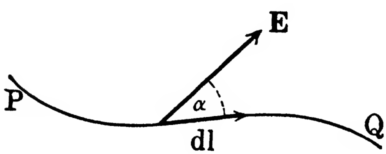


FIG. 79

There are several types of force, to any one or to any combination of which E may be due. If $E = E_q$ represents electrostatic forces only, with which we are already familiar, the integral of (48-3) is identical with the potential drop from P to Q as defined in article 7 and is therefore independent of the path. With other types of force, which we shall study later, there is in general no definite potential difference between the two points, that is, the integral is not independent of the path. We must therefore specify the e.m.f. from P to Q along a given path as the work done on the unit charge when it is carried from P to Q by that path. We are usually interested in the electromotive force around a complete circuit. Since the integral of $E_q \cos \alpha dl$ along any path is the potential difference $V_P - V_Q$ of the end

points, its value is zero when these points are brought together to form a circuit. Analytically,

$$\mathcal{E}_q = \oint E_q \cos \alpha dl = 0, \tag{48-4}$$

where the symbol $\oint$ indicates integration around a closed path. It is therefore clear that *no static distribution of charge can produce an e.m.f. around a circuit*, a fact important in circuit theory. In general, however, the electromotive force

$$\mathcal{E} = \oint E \cos \alpha dl \tag{48-5}$$

around a circuit does not vanish, but depends on the path followed.

We have previously defined the electrostatic unit of charge, and the practical unit or coulomb. Each of these gives rise to a unit of current, although current is rarely measured in e.s.u. The practical unit, that is, a coulomb per second, is called the *ampere*. In addition to the two units just mentioned there is a third unit of current whose existence is due to the fact that a current produces a magnetic field, a phenomenon to be studied in detail in Chapter VII. Thus if a wire carrying a current is bent into a circle a magnetic force is developed at the center of the circle, perpendicular to its plane. If the circle is 1 centimeter in radius and the magnetic force at the center is 2π gauss, we have by definition an *electromagnetic unit of current* flowing. Its magnitude is equal to that of ten amperes.

On account of its magnetic field a coil of wire carrying a current exerts a torque on a magnet, and *vice versa*. If one is fixed and the other suspended so that it may rotate we have a means of detecting and measuring current. A current indicator of this type is called a *galvanometer*. A complete description of such instruments is deferred until a study of electromagnetism is made. Some important applications of the galvanometer are given, however, in connection with various electrical measurements described in this chapter.

49. Metallic Conduction and Ohm's Law. — Let us now examine the mechanism of conduction in solids, that is, in metals, as all good conductors are metals. We picture the interior of a conductor as a three dimensional lattice of atoms with free electrons moving about between the atoms and frequently colliding with them. As the effective radius of an atom is of the order of $(10)^{-8}$ centimeter, while that of an electron is of the order of $(10)^{-13}$ centimeter, we may neglect the collisions between electrons in comparison with the collisions between electrons and atoms. The atoms vibrate about their equilibrium positions, and the electrons move, with a mean energy determined by the absolute temperature. In fact, the atoms have the usual thermal agitation of any solid, while the electrons collectively behave like a gas at the temperature of the conductor.

As long as no external field is applied to the conductor the average field in the interior is zero, since there is as much positive charge as negative in any small volume. Therefore there is on the whole no motion of free charges through the conductor. If now we apply a field of intensity E each charge experiences an acceleration of magnitude eE/m, where m is the mass of the charge. This superposes on the random thermal motions of the charges a general drift, which constitutes the current. The drift velocity is small compared to the thermal velocities, so that the two motions may be treated as independent. The *mean free time* t between successive collisions of a free charge with some atom is thus determined by the structure of the conductor and the temperature, but not by E.

We may now calculate the mean drift velocity and hence the current. Between collisions the drift velocity is increased on the average by the amount $(eE/m)t$. The effect of each collision however is to restore the random thermal distribution of velocities, that is, to reduce the drift velocity to zero. Therefore the mean drift velocity v is given by

$$v = \frac{0 + \left(\frac{eE}{m}\right)t}{2} = \frac{1}{2}\left(\frac{eE}{m}\right)t,$$

and

$$j = \left(\frac{ne^2t}{2m}\right)E. \tag{49–1}$$

Defining the *electrical conductivity* σ as the current density produced by a field of unit strength we have

$$\sigma = \frac{j}{E} = \frac{ne^2t}{2m}. \tag{49–2}$$

Under all ordinary conditions the conductivity is a characteristic of the conducting substance independent of j or E. It does, however, vary with the temperature. According to the kinetic theory of gases t, and hence σ, is inversely proportional to the square root of the absolute temperature. Experimentally it is found that σ is more nearly proportional to the inverse first power of the temperature. This discrepancy is not surprising in view of the approximate method used in deducing (49–2).

A more satisfactory check on the validity of the electron theory of conduction is obtained by considering thermal conduction as well as electrical. The mean energy of both the electrons and the atoms in a body depends on the temperature, so that when a temperature gradient exists, there is also an energy gradient. Energy is transferred from a region of higher temperature to one of lower by diffusion of the electrons. We may ascribe heat conduction almost entirely to this cause, since in dielectrics where there are no free charges, there is very small conduction of heat. By an analysis similar to that used for electrical conductivity it may be shown * that the thermal conductivity K is given by

$$K = \left(\frac{3nk^2t}{2m}\right)T, \tag{49–3}$$

where k is a universal constant of magnitude $1.37(10)^{-16}$ erg per degree, and T is the absolute temperature.

Taking the ratio of (49–3) to (49–2) we find

$$\frac{K}{\sigma} = 3\left(\frac{k}{e}\right)^2 T, \tag{49–4}$$

* Page: Introduction to Theoretical Physics, p. 391.

a relation known as the *law of Wiedemann and Franz.* It is found to be in good agreement with experiment, at least in the case of the best conductors, such as gold, silver and copper.

Returning now to (49–2) the *resistivity* ρ is defined as the reciprocal of the conductivity, so that, using (48–2) to introduce the total current,

$$E = \rho j = \frac{\rho}{A} i. \tag{49–5}$$

Let us apply the last equation to a wire of length l, the composition and cross-section being uniform throughout its length. Since E is then constant and directed along the wire, El is the electromotive force and we may write

$$\mathcal{E} = \frac{\rho l}{A} i. \tag{49–6}$$

Now since $\rho = 1/\sigma$ the quantity $\rho l/A$ depends only on the absolute temperature under ordinary conditions, being in fact approximately proportional to it. If we denote $\rho l/A$ by R we have

$$\mathcal{E} = Ri, \tag{49–7}$$

which is *Ohm's law.* R is called the *resistance* of the conductor. In the practical system of units, where $\mathcal{E}$ is measured in volts and i in amperes, R is measured in *ohms.* Very large resistances are sometimes measured in *megohms*, a megohm being a million ohms.

Ohm's law may be stated in general terms as follows: *The ratio of the electromotive force between two points on a conductor to the current flowing between these points is a constant, at any given temperature, known as the resistance.*

The passage of current through a conductor is evidently attended by an evolution of heat, since the moving charges lose their energy to the atoms at each collision. The heat generated per second in a conductor is easily calculated. Using the same notation as above, the work done per unit length of conductor in a time dt is

$$AneEvdt = Eidt,$$

by (48–1). The work for the entire conductor is

$$(\int E dl) i dt = \mathcal{E} i dt,$$

so that the work done per second, called the *power*, is $\mathcal{E}i$. This energy all appears in heat, as there is no storage of energy in the interior of the conductor. Denoting power by $\mathcal{P}$, and using (49–7),

$$\mathcal{P} = \mathcal{E}i = Ri^2. \tag{49–8}$$

In e.s.u. or e.m.u. power is measured in ergs per second. In the practical system the unit is a *joule per second* or *watt*. As a joule is $(10)^7$ ergs the practical unit of power is $(10)^7$ times as great as the c.g.s. unit. The number of calories developed per second is obtained by multiplying the power in watts by 0.238.

A table of resistivities is appended for reference. The metals are arranged in order of increasing resistivity, the values being given in ohm centimeters for a temperature of 0° C. Note that the resistivity is the resistance of a unit cube of the given material. The total resistance of a conductor of any length and cross-section is calculated by means of the relation $R = \rho l/A$. The resistivity at a temperature not greatly different from 0° C is given by the formula $\rho = \rho_0 (1 + \alpha t)$ where ρ_0 is the resistivity at 0°C, t is the centigrade temperature, and α is the *temperature coefficient*.

Substance	ρ_0 (ohm cm)	α (1/deg C)
Metals:		
Silver	$1.53(10)^{-6}$	$40(10)^{-4}$
Copper (annealed)	1.65	43
Gold	2.25	40
Aluminum	2.88	38
Tungsten	4.44	51
Tin	10.4	45
Platinum	10.3	38
Iron	10.8	62
Lead	19.3	43
Mercury	94.1	9
Alloys:		
Brass	6.50	10
Manganin	42.0	0
Nichrome	110.0	2

50. Conversion of Units. — Although electrostatic formulas are usually deduced in e.s.u., and magnetic formulas in e.m.u., it is sometimes desirable to convert from one system of units to the other. Moreover, laboratory measurements on either type of quantity are always made in practical units. Transformations are simply effected when it is remembered that any symbol in a formula represents the measure of a definite physical entity in terms of a given unit magnitude. Thus, suppose we observe a certain current. Let its measure in e.s.u. be i_s, in e.m.u., i_m, and in practical units, i_p. Now the magnitude of the e.m.u. of current is ten times as great as that of the practical unit. Therefore the measure of the current in e.m.u. is one-tenth its measure in amperes, so that the equation $i_p = 10 i_m$ is a numerical identity which may be substituted in any formula. Similarly the ampere is $3(10)^9$ times as great as the e.s.u. of current, so that the proper substitution equation is $i_s = 3(10)^9 i_p$. These and several other common conversion relations are included in the table below. A complete conversion table is to be found at the beginning of the book.

Quantity	Conversion Equations
Charge.................	$Q_m = \dfrac{1}{3(10)^{10}} Q_s = \dfrac{1}{10} Q_p$
Potential and e.m.f.......	$V_m = 3(10)^{10} V_s = (10)^8 V_p$
Electric intensity........	$E_m = 3(10)^{10} E_s = (10)^8 E_p$
Capacity...............	$C_m = \dfrac{1}{9(10)^{20}} C_s = \dfrac{1}{(10)^9} C_p$
Current................	$i_m = \dfrac{1}{3(10)^{10}} i_s = \dfrac{1}{10} i_p$
Resistance..............	$R_m = 9(10)^{20} R_s = (10)^9 R_p$

To illustrate the use of the table let us transform (18–1), $Q_s = C_s V_s$. Making the indicated substitutions,

$$3(10)^{10} Q_m = 9(10)^{20} C_m \frac{V_m}{3(10)^{10}}, \qquad Q_m = C_m V_m;$$

and

$$3(10)^9 Q_p = 9(10)^{11} C_p \frac{V_p}{300}, \qquad Q_p = C_p V_p.$$

Hence this equation takes exactly the same form in all three systems of units. The same thing is true of Ohm's law, and a few other relations, but in general numerical factors appear. For example, (18–5), $C_s = \kappa ab/(b - a)$, transforms to

$$C_m = \frac{1}{9(10)^{20}} \frac{\kappa ab}{b - a},$$

$$C_p = \frac{1}{9(10)^{11}} \frac{\kappa ab}{b - a}.$$

A further discussion of units appears in Chapter XII.

51. The Voltaic Cell. — In order to maintain a steady flow of current a constant electromotive force is necessary. This is readily obtained by chemical means, for if two dissimilar conductors are dipped in a conducting liquid which reacts with them there is in general a potential difference between them, which is maintained even when current flows steadily from one to the other through some external resistance. Such a device is called a *voltaic cell.* The two conductors are the *poles*, the positively charged one being the *positive pole*, and the other the *negative pole.*

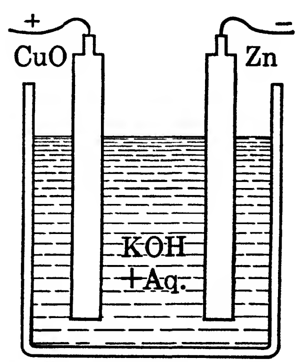


Fig. 80

Let us consider for example copper oxide and zinc dipped in an aqueous solution of potassium hydroxide (Fig. 80). This is an *Edison cell.* Within the cell chemical forces exist which tend to transfer charge from one pole to the other. These forces are capable of doing a definite amount of work $\mathcal{E}_c$ per unit charge. They build up a positive charge on one pole, the CuO, and a negative charge on the other, giving rise to electrostatic forces which, in the interior of the cell, act in opposition to the chemical forces. Let E_c be the force per unit charge in the cell due to chemical causes, and E_q the force per unit charge due to static

distribution of charge. Then the electromotive force around a circuit consisting of any path from the positive pole to the negative outside the cell and any path from the negative to the positive inside—regardless of whether current is flowing or not—is given by

$$\mathcal{E} = \int_o E_q \cos \alpha dl + \int_i (E_q + E_c) \cos \alpha dl$$
$$= \oint E_q \cos \alpha dl + \int_i E_c \cos \alpha dl, \qquad (51\text{–}1)$$

where the subscripts o and i indicate integration along the outside path and the inside path respectively, and $\oint E_q \cos \alpha dl$ is the line integral of the electrostatic force around the complete circuit. By (48–4) this line integral is zero, and, as the last integral in (51–1) is $\mathcal{E}_c$, we have

$$\mathcal{E} = \mathcal{E}_c. \qquad (51\text{–}2)$$

That is, *the e.m.f. of the circuit equals the chemical e.m.f. of the cell.*

First suppose that the poles are not connected externally. Then no current flows, and the condition of equilibrium in the cell requires $E_q + E_c = 0$ everywhere inside. Hence, using (51–2) and (51–1),

$$\mathcal{E}_c = \mathcal{E} = \int_o E_q \cos \alpha dl. \qquad (51\text{–}3)$$

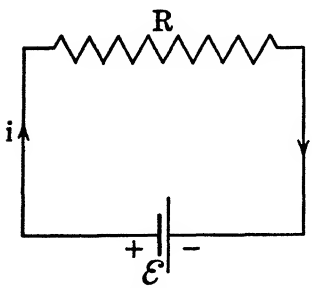


Fig. 81

The integral represents the potential difference of the poles; so that *the chemical e.m.f. of the cell equals the open circuit potential difference.*

Next let us connect the poles by a conductor of resistance R. A current i flows, the circuit being shown schematically in Fig. 81. From (51–2) we know the total work done on a unit charge as it passes around the circuit. While this work is *generated* chemically in the cell, it can be *performed* only in overcoming resistance.

If we denote the *internal resistance* of the cell by r, the total work done per unit charge is, by Ohm's law, $(R + r)i$, so that

$$\mathcal{E} = (R + r)i, \qquad i = \frac{\mathcal{E}}{R + r}. \tag{51-4}$$

The quantity r is not strictly an ohmic resistance, being dependent to some extent on the current and on the past history of the cell. In the type of cell described above it is usually negligibly small, but in some cells, particularly *dry cells* in which the liquid is replaced by a paste, the internal resistance may be appreciable. Observe that the potential difference of the poles, which is given by $Ri = \mathcal{E} - ri$, is now less than its open circuit value by an amount ri.

It is interesting to note that while the charge on the poles does not contribute to the e.m.f. of the circuit, it provides the mechanism by which work generated internally is performed externally. In fact, if r is negligible,

$$\mathcal{E} = Ri = \int_o E_q \cos \alpha dl,$$

and (51–1) reduces to

$$0 = \int_i (E_q + E_c) \cos \alpha dl,$$

which indicates that the internal chemical forces are everywhere neutralized by the electrostatic forces, so there is no work done on a unit charge as it passes through the cell. On the other hand if r exists,

$$\mathcal{E} - ri = Ri = \int_o E_q \cos \alpha dl,$$

and (51–1) becomes

$$ri = \int_i (E_q + E_c) \cos \alpha dl.$$

There is now an amount of work ri performed per unit charge in the cell. But even in this case there is a transfer of work; for electrochemical evidence indicates that the source of $\mathcal{E}_c$ lies in

the contact between the liquid or *electrolyte*, and the poles, while the work ri is performed throughout the electrolyte.

Referring to $\mathcal{E}$ as the *applied e.m.f.*, we may state Ohm's law for a complete circuit in which a steady or *direct* current is flowing in the form: *The current equals the applied e.m.f. divided by the total resistance of the circuit.*

The e.m.f. of all commonly used cells lies between one and two volts. When a higher e.m.f. is desirable it may be obtained by connecting several cells in *series*, that is, by connecting the positive pole of one to the negative pole of the next (Fig. 82*a*). As the work done in carrying a unit charge around a circuit passing through all the cells is the sum of the $\mathcal{E}_c$'s, the resultant e.m.f. is the sum of the separate e.m.f.'s. Obviously a *parallel* connection (Fig. 82*b*) of identical cells does not change the e.m.f. although it increases the current capacity. Any combination of cells is called a *battery*.

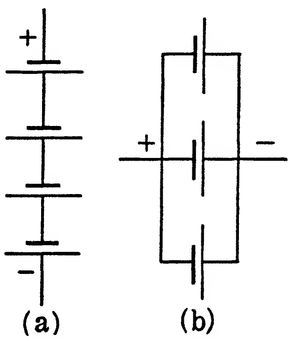


FIG. 82

Problem 51a. Show that if n identical cells of e.m.f. $\mathcal{E}$ and resistance r are connected in series they will produce a current through an external resistance R equal to $n\mathcal{E}/(R + nr)$, while if they are connected in parallel the current is $n\mathcal{E}/(nR + r)$.

Problem 51b. Using the results of the preceding problem find the relation between R and r when (*a*) the series connection produces a greater current than the parallel, (*b*) the parallel produces greater than the series, (*c*) both produce the same. Ans. (*a*) $R > r$, (*b*) $R < r$, (*c*) $R = r$.

52. Combinations of Resistances and Kirchhoff's Laws. — In the construction of circuits it is often necessary to connect several resistances together, in *series* (Fig. 83*a*), or in *parallel* (Fig. 83*b*), or in more complicated ways to form *networks*.

In the series case the total e.m.f. $\mathcal{E}$ of the combination is clearly the sum of the individual e.m.f.'s of the resistance

elements. If a total current i is flowing we have

$$\mathcal{E} = R_1 i + R_2 i + \cdots R_k i = \left[\sum_1^k R_j \right] i. \qquad (52\text{–}1)$$

Now the equivalent or total resistance R of the series combination is given by $\mathcal{E} = Ri$. Comparing this with (52–1) we see that

$$R = \sum_1^k R_j. \qquad (52\text{–}2)$$

In the parallel case the e.m.f. is the same for all the resistance elements, while the currents $i_1, i_2, \cdots i_k$ are in general different. Applying Ohm's law to each resistance,

$$\mathcal{E} = R_1 i_1 = R_2 i_2 = \cdots R_k i_k, \qquad (52\text{–}3)$$

and also

$$\mathcal{E} = Ri, \qquad (52\text{–}4)$$

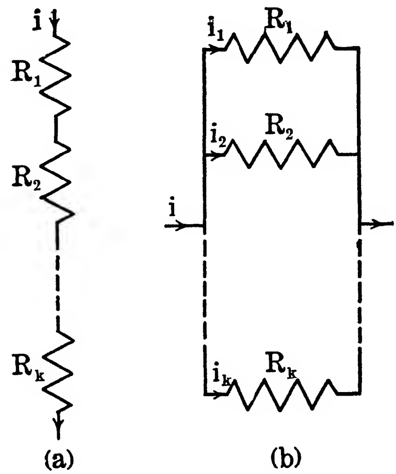


FIG. 83

where R and i denote total resistance and total current as before. As $i = i_1 + i_2 + \cdots i_k$ we have with the aid of (52–3) and (52–4),

$$\frac{\mathcal{E}}{R} = \frac{\mathcal{E}}{R_1} + \frac{\mathcal{E}}{R_2} + \cdots \frac{\mathcal{E}}{R_k} = \mathcal{E} \sum_1^k \frac{1}{R_j},$$

or

$$\frac{1}{R} = \sum_1^k \frac{1}{R_j}. \qquad (52\text{–}5)$$

In either of the cases considered above, the calculation of R is but a means to an end, of course. In most circuit problems we are interested in the currents in the various elements of the circuit, together with the corresponding e.m.f.'s. Occasionally only certain ratios of these quantities are required.

Let us now investigate resistance networks in general. Although it sometimes happens that a network may be analyzed

into series and parallel groups, more often — in fact whenever a branch point in the circuit is connected to more than two other branch points — no such simple treatment is possible. The

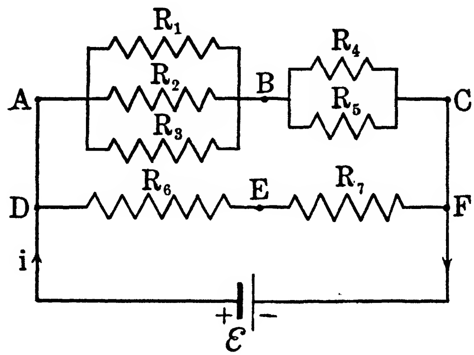


FIG. 84

first case is illustrated in Fig. 84. The resistance from A to B and that from B to C are obtained by use of (52–5). These are combined by (52–2) to obtain the resistance of the path ABC.

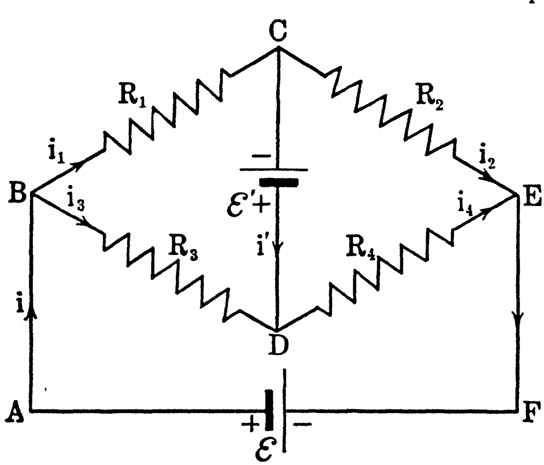


FIG. 85

Similarly the resistance of DEF is calculated and combined with that of ABC by (52–5) to obtain the total resistance of the circuit. An example of the second and more complex case is

shown in Fig. 85. To analyze this circuit we must evidently return to fundamentals, where two facts at once present themselves. First, there can be no accumulation of charge at a branch point, so that the sum of the currents flowing to it must equal the sum of the currents flowing from it. Second, if we take a unit positive charge around any closed path in the network, the total work done on the charge must equal the sum of the applied e.m.f.'s in the given path. Having arbitrarily chosen positive directions, as indicated in the figure, we may put the above statements in a more analytical form, known as *Kirchhoff's laws*.

Law 1. The algebraic sum of all the currents meeting at a point is zero.

Law 2. The algebraic sum of the Ri terms around any closed path equals the algebraic sum of the applied e.m.f.'s in the given path.

Let us now find the six unknown currents of Fig. 85, in terms of the R's and $\mathcal{E}$'s, as an illustration of the application of Kirchhoff's laws to a circuit problem. We require six independent equations, that is, six equations no one of which can be derived from a combination of the others. From *Law 1* we may write

$$\left.\begin{aligned} i - i_1 - i_3 &= 0, & (1)\\ i_1 - i_2 - i' &= 0, & (2)\\ i_3 + i' - i_4 &= 0, & (3)\end{aligned}\right\}(52\text{–}6)$$

and from *Law 2*

$$\left.\begin{aligned} R_1i_1 + R_2i_2 - R_4i_4 - R_3i_3 &= 0, & (4)\\ R_1i_1 - R_3i_3 &= \mathcal{E}', & (5)\\ R_1i_1 + R_2i_2 &= \mathcal{E}. & (6)\end{aligned}\right\}(52\text{–}7)$$

There are several other equations directly obtainable, but they are not independent of the six chosen. For example, the equation $i_2 + i_4 - i = 0$, applying to the branch point E, may be obtained by adding the first three equations above. Evidently it may be substituted for any of these, so that the group of equations to be solved is to some extent arbitrary.

The algebraic solution of the equations (52–6) and (52–7), which is most easily effected by means of determinants, is left to the reader. If we set

$$\Delta \equiv R_1R_2R_3 + R_2R_3R_4 + R_3R_4R_1 + R_4R_1R_2,$$

we find for the solution

$$\left.\begin{aligned}
i_1 &= \frac{R_3(R_2 + R_4)\mathcal{E} + R_2(R_3 + R_4)\mathcal{E}'}{\Delta},\\
i_2 &= \frac{R_4(R_1 + R_3)\mathcal{E} - R_1(R_3 + R_4)\mathcal{E}'}{\Delta},\\
i_3 &= \frac{R_1(R_2 + R_4)\mathcal{E} - R_4(R_1 + R_2)\mathcal{E}'}{\Delta},\\
i_4 &= \frac{R_2(R_1 + R_3)\mathcal{E} + R_3(R_1 + R_2)\mathcal{E}'}{\Delta},\\
i &= \frac{(R_1 + R_3)(R_2 + R_4)\mathcal{E} + (R_2R_3 - R_4R_1)\mathcal{E}'}{\Delta},\\
i' &= \frac{(R_2R_3 - R_4R_1)\mathcal{E} + (R_1 + R_2)(R_3 + R_4)\mathcal{E}'}{\Delta}.
\end{aligned}\right\}\quad (52\text{–}8)$$

Observe that under certain conditions some of the currents may be negative. A negative current is, of course, one which flows in a direction opposite to that arbitrarily chosen as positive.

There are several points of interest in connection with (52–8). In the first place each e.m.f. makes its contribution to the various currents independently of the other e.m.f. This is a general property of networks, and often a useful one; for a complicated network containing a number of e.m.f.'s may be solved for each separately, and the complete solution obtained by addition. Again, we see that, when $\mathcal{E}' = 0$, i' has the same form as i when $\mathcal{E} = 0$. This illustrates another important general property of networks; namely, if an e.m.f. $\mathcal{E}$ in the jth branch of a network produces a current i in the kth branch, then an e.m.f. $\mathcal{E}$ in the kth branch will produce the same current i in the jth branch. The foregoing statement is sometimes called the *reciprocity theorem.* Finally, note that if $R_2R_3 = R_4R_1$, $\mathcal{E}$ makes no contribution to i', and $\mathcal{E}'$ makes none to i. Any two branches of a network with this mutual property are said to be *conjugate.*

Problem 52a. Resistances of 3, 5 and 7 ohms respectively are connected in parallel, the entire group being placed in series with a resistance of 1 ohm. What is the total resistance? Ans. 2.48 ohm.

Problem 52b. The combination of resistances described in the preceding problem is connected to a battery of two identical cells in series. Each cell has an e.m.f. of 1.5 volts and an internal resistance of 0.1 ohm. Find the total current which passes through the battery, and also the partial currents in the 3, 5 and 7 ohm resistances. Ans. 1.12 amp; 0.55, 0.33, 0.24 amp.

Problem 52c. Eight equal resistances R are arranged to form a square with diagonals connected at the midpoint. Find the total resistance between adjacent corners of the square; between opposite corners. Ans. $\frac{8}{15}R$, $\frac{2}{3}R$.

Problem 52d. Twelve equal resistances R are arranged to form a hexagon with diagonals connected at the midpoint. Find the total resistance between adjacent vertices; between opposite vertices. Ans. $\frac{11}{20}R$, $\frac{4}{5}R$.

Problem 52e. N identical cells of e.m.f. $\mathcal{E}$ and internal resistance r are connected in p rows each of which contains s cells in series. The p rows are connected in parallel with an external resistance R. Find the current through R and show that it is greatest when the cells are so grouped that the total resistance of the cells is as nearly equal as possible to R.

$$\text{Ans. } \frac{N\mathcal{E}}{pR + sr}.$$

Problem 52f. Two batteries of electromotive forces 6 and 10 volts and resistances 0.5 and 1 ohm respectively are connected in parallel with a resistance of 12 ohms. Find the current through each branch of the circuit and the potential difference between the two junctions. Ans. $-$ 2.27 amp, 2.86 amp, 0.59 amp; 7.14 volt.

Problem 52g. By means of Kirchhoff's laws show that if a resistance R' is placed in series with $\mathcal{E}'$ (Fig. 85) the current through it is given by

$$i' = \frac{i_0'}{1 + \frac{R'}{\Delta}(R_1 + R_2)(R_3 + R_4)},$$

where i_0' is the current through $\mathcal{E}'$ when $R' = 0$ and Δ has the same significance as in (52–8).

53. The Wheatstone Bridge. — In laboratory practice rapid and accurate measurement of resistance is essential. A variety of methods has been devised, almost all of which depend upon

a special type of resistance network known as a *Wheatstone bridge* (Fig. 86). This differs from the network of Fig. 85 only in that $\mathcal{E}'$ is replaced by a galvanometer or other current indicator G, of resistance R_g. Evidently i_g may be obtained from the

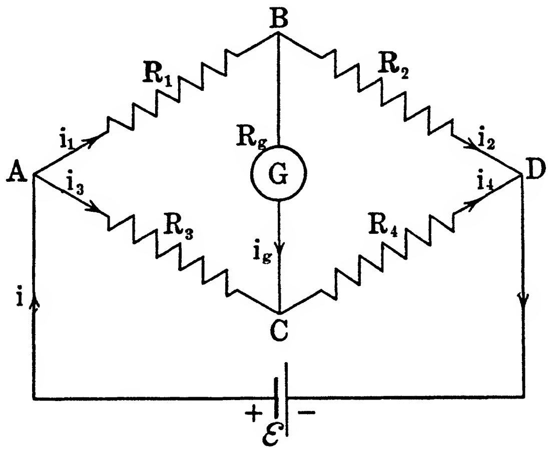


FIG. 86

last equation of (52–8) by substituting $-R_g i'$ for $\mathcal{E}'$, and solving for i' ($\equiv i_g$). This gives

$$i_g = \frac{(R_2R_3 - R_4R_1)\mathcal{E}}{\Delta + R_g(R_1 + R_2)(R_3 + R_4)}, \tag{53–1}$$

and we see that $i_g = 0$ when $R_2R_3 = R_4R_1$. As a matter of fact the condition for $i_g = 0$ may be found without solving the network equations. For if $i_g = 0$ there can be no difference of potential between B and C which requires that the potential drop from A to B equals that from A to C. That is,

$$R_1i_1 = R_3i_3. \tag{53–2}$$

Similarly we must have

$$R_2i_2 = R_4i_4. \tag{53–3}$$

Since with $i_g = 0$ we have $i_1 = i_2$ and $i_3 = i_4$, we may divide

(53–2) by (53–3) and obtain

$$\frac{R_1}{R_2} = \frac{R_3}{R_4}, \tag{53–4}$$

which is the same relation as found from (53–1).

Now let us suppose that one of the resistances, say R_1, is unknown, and that R_2, R_3, R_4 are known, at least one of them being variable. The variable resistance is adjusted until G indicates that $i_g = 0$. Then the relation (53–4) is established between the resistances and R_1 becomes known in terms of R_2, R_3 and R_4. The adjustment of the bridge to the condition of no current in the galvanometer arm is called *balancing*, and is the essential feature of resistance measurements by the bridge method. Calibration of G is not required, since it is used only as a null indicator.

Although the existence of the balance depends only upon (53–4), the *accuracy* with which the balance can be established experimentally is dependent on the characteristics of G and on the magnitudes of the resistances and of the applied e.m.f. In order to discover the best arrangement of the bridge let us suppose balance is obtained by adjustment of R_3. Then, since the precision of adjustment is greater, the larger the galvanometer deflection for a given fractional change of R_3 at the point of balance, we may define the *sensitivity* S of the bridge by

$$S = CR_3\left(\frac{\partial i_g}{\partial R_3}\right)_0, \tag{53–5}$$

where the subscript 0 indicates that the derivative is to be evaluated at the balance point, and C is the galvanometer deflection per unit current. Using (53–1) we find

$$\begin{aligned} S &= CR_3\left\{\frac{R_2}{\Delta + R_g(R_1 + R_2)(R_3 + R_4)}\right. \\ &\qquad \left. - \frac{(R_2R_3 - R_4R_1)[R_1R_2 + R_2R_4 + R_4R_1 + R_g(R_1 + R_2)]}{[\Delta + R_g(R_1 + R_2)(R_3 + R_4)]^2}\right\}_0 \mathcal{E} \\ &= C\frac{R_2R_3\mathcal{E}}{\Delta_0 + R_g(R_1 + R_2)(R_3 + R_4)} \\ &= C\frac{R_4R_1\mathcal{E}}{\Delta_0 + R_g(R_1 + R_2)(R_3 + R_4)}. \end{aligned} \tag{53–6}$$

Evidently S has exactly the same form if adjustment is made by means of R_2 or R_4, and therefore for a given set of resistances the bridge is equally sensitive to a variation of any arm. If all the known resistances are variable, as is often the case, adjustment to balance should be made by means of that resistance which can be varied by the smallest fraction of itself.

We first investigate the condition for a maximum value of S, with a given battery and galvanometer. That is, $\mathcal{E}$ and R_g are constant, but R_2, R_3 and R_4 may be varied, subject to the relation $R_2R_3 = R_4R_1$. In addition to this mathematical restriction there are certain physical limitations which must be observed, for each resistance element has a definite *current capacity* or maximum safe current, the magnitude of which depends on the construction of the element. Therefore with the given e.m.f. there is a minimum permissible value for $R_1 + R_2$, and also for $R_3 + R_4$. The minimum value may be different in the two cases, of course. In fact, the current capacity of a given type of resistance element or of a variable resistance is usually greater for low resistance than for high. We may avoid any restriction on R_2 by supposing that $\mathcal{E}$ is small enough so that the current through R_1 can never be excessive regardless of the value of R_2. We cannot, however, escape the condition

$$R_3 + R_4 = R + \delta, \qquad \delta \geq 0, \tag{53-7}$$

where R is the minimum safe value of $R_3 + R_4$. Since (53–6) reduces to the simple form

$$S = \frac{C\mathcal{E}}{(R_1 + R_2 + R_3 + R_4) + R_g\left(\dfrac{R_1}{R_2} + 2 + \dfrac{R_2}{R_1}\right)}, \tag{53-8}$$

the desired maximum corresponds to a minimum of the denominator

$$D = (R_1 + R_2 + R_3 + R_4) + R_g\left(\frac{R_1}{R_2} + 2 + \frac{R_2}{R_1}\right), \tag{53-9}$$

subject to the conditions

$$\left.\begin{aligned} R_2R_3 &= R_4R_1, \\ R_3 + R_4 &= R + \delta, \qquad \delta \geq 0. \end{aligned}\right\} \tag{53-10}$$

Minimum D evidently requires $\delta = 0$, and in addition $\dfrac{\partial D}{\partial R_2} = 0$. That is,

$$\frac{\partial}{\partial R_2}\left\{(R_1 + R_2 + R) + R_g\left(\frac{R_1}{R_2} + 2 + \frac{R_2}{R_1}\right)\right\} = 0,$$

which leads to

$$1 - R_g\frac{R_1}{R_2^2} + R_g\frac{1}{R_1} = 0,$$

or

$$R_2 = R_1\sqrt{\frac{R_g}{R_1 + R_g}}. \tag{53-11}$$

From (53–10) and (53–11) with $\delta = 0$ we find

$$\left.\begin{aligned} R_3 &= \frac{R}{1 + \sqrt{\dfrac{R_g}{R_1 + R_g}}} \\ R_4 &= \frac{R\sqrt{\dfrac{R_g}{R_1 + R_g}}}{1 + \sqrt{\dfrac{R_g}{R_1 + R_g}}}. \end{aligned}\right\} \tag{53-12}$$

Inspection of these results shows that if R_g is large compared to R_1, the best arrangement of the bridge is $R_1 = R_2$ and $R_3 = R_4 = R/2$. When there is a choice between several galvanometers with different values of R_g, differences in C must also be taken into account, maximum S for each galvanometer being calculated with the aid of (53–8). If, however, a galvanometer is designed especially for a given bridge, R_g being adjusted to any desired value by proper choice of wire size in the winding, we have, approximately at least, $C = c\sqrt{R_g}$. That this is so may be seen by considering two galvanometer coils of the same dimensions, one coil having half as many turns of twice as large wire as the other. The current sensitivities will be in the ratio of one to two, while the resistances will be in the ratio of one to four. Substituting $C = c\sqrt{R_g}$ in (53–8) and

setting $\dfrac{\partial S}{\partial R_g}$ equal to zero gives

$$R_g = R_2R_3\frac{R_1 + R_2 + R_3 + R_4}{(R_1 + R_2)(R_3 + R_4)} = \frac{(R_1 + R_3)(R_2 + R_4)}{R_1 + R_2 + R_3 + R_4},$$

or

$$\frac{1}{R_g} = \frac{1}{R_1 + R_3} + \frac{1}{R_2 + R_4}. \tag{53-13}$$

Thus *the resistance of the galvanometer should equal the parallel resistance of the bridge.* This is a perfectly general result, quite independent of other conditions or restrictions.

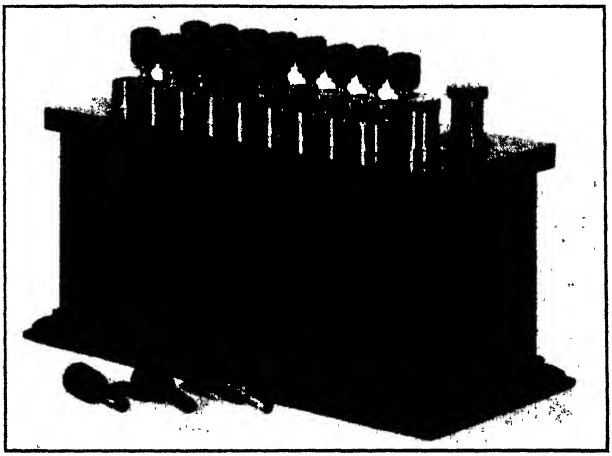
It only remains to solve (53–11) to (53–13) inclusive for R_2, R_3, R_4 and R_g in terms of R_1 and R. In practice we usually find that R is small compared with R_1. Then approximately $R_g = R_2R_3/R = R_4R_1/R$ and the solution is

$$\left.\begin{aligned} R_2 &= \tfrac{1}{2}R_1, & R_g &= \tfrac{1}{3}R_1, \\ R_3 &= \tfrac{2}{3}R, & R_4 &= \tfrac{1}{3}R. \end{aligned}\right\} \tag{53-14}$$

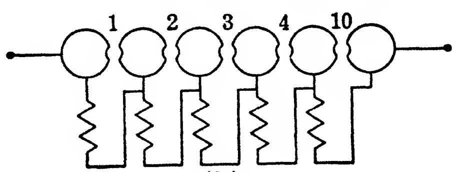
This optimum arrangement of the bridge can rarely be attained in the laboratory due to limitations in equipment of one sort or another. However, the equations (53–14) may serve as a guide, at least, to the best arrangement under the existing experimental limitations. Evidently for a given set of resistances the galvanometer should be connected between the junction of the two larger resistances and the junction of the two smaller resistances. Also, the galvanometer resistance should correspond to the resistance to be measured. That is, while the standard galvanometer resistance of about 100 ohms is quite satisfactory for ordinary measurements, a bridge designed to measure a few ohms should be supplied with a low resistance galvanometer, and a bridge for high resistance measurements should have a high resistance galvanometer. The application of these principles is made clear in the following articles.

54. Resistance Boxes and Resistance Standards. — A variable resistance is usually enclosed in a box, with some external device for adjusting its magnitude, the entire assembly being

called a *resistance box*. The most common types are the *plug box* (Fig. 87) and the *dial box* (Fig. 88). The plug box consists of a number of resistance elements of different magnitudes connected in series. Each junction between elements is connected to a metal block; so that if a metal plug is inserted between adjacent blocks the corresponding resistance is short-circuited, that is, removed from the circuit (Fig. 87*b*). The total resistance



(a)



(b)

Fig. 87

of the box is thus the sum of the resistances whose plugs have been removed. The resistance elements are spools of some resistance wire with a small temperature coefficient, such as manganin wire (See Table, p. 165). The lower resistance elements are usually wound with larger wire than the higher resistances, so that their current capacity is greater, as mentioned in article 53.

The dial box also consists of a number of resistance elements in series, the junctions in this case being connected to a set of metal buttons over which a contact arm A moves. This *dial*

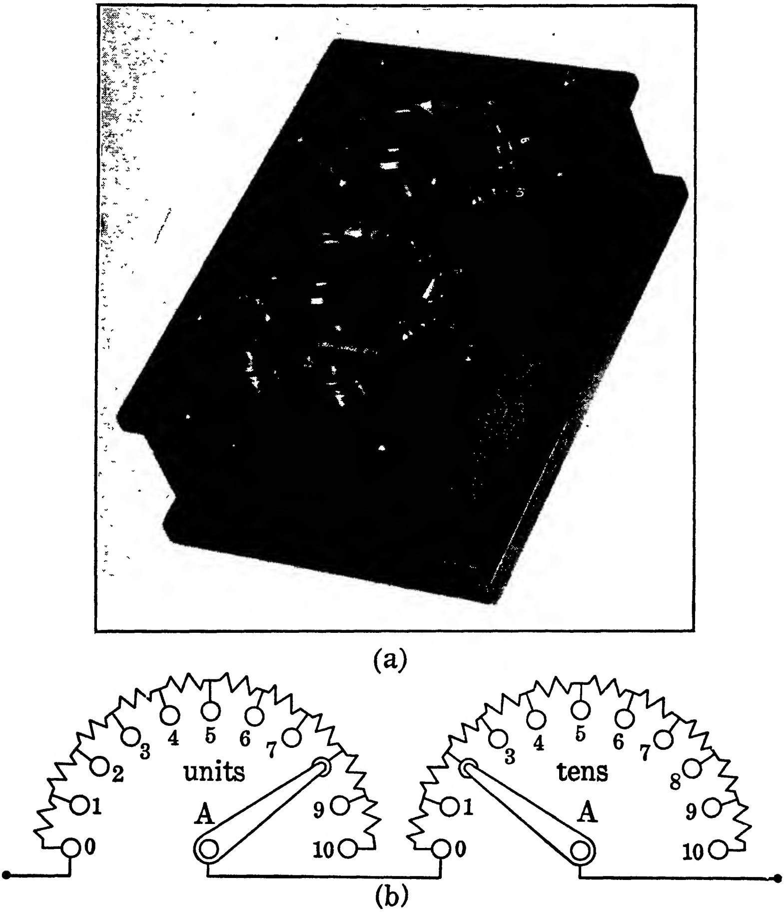


FIG. 88

switch allows any number of the elements to be included in the circuit. Usually the elements are arranged in groups of ten of equal magnitude, successive groups however differing in magnitude by a factor of ten (Fig. 88*b*). With this scheme the posi-

tion of the switch arms indicates the total resistance directly, a convenience not offered by the plug box. The latter has some advantages, however, as it is less bulky than the other and less subject to errors arising from contact resistance, the plugs being more positive in action than the sliding contact arm.

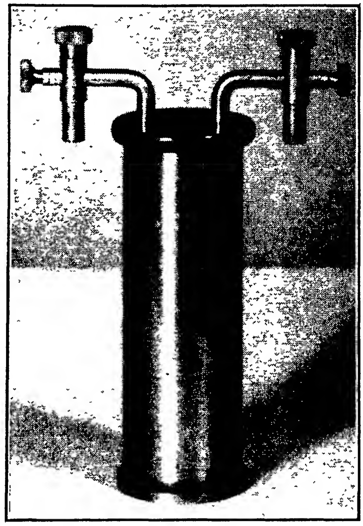


FIG. 89

For the most accurate resistance determinations *fixed* standards are used. In the Bureau of Standards type (Fig. 89) the resistance coil is sealed into an oil filled container. A thermometer well in the center allows the temperature to be determined. The terminals are heavy lugs of such shape that they may be dipped in mercury cups to avoid contact resistance.

A special form of known resistance called a *shunt* is of great utility. It is often desirable to reduce the effective sensitivity of a current indicator in order to increase its range or to prevent excessive deflections. A small resistance connected in parallel with the instrument achieves this result by shunting or by-passing all but a small fraction of the total current. Suppose an instrument of resistance R_g is shunted by a resistance αR_g. If the total current through the combination is i and that through the instrument only is i_g we have

$$R_g i_g = \frac{1}{\dfrac{1}{R_g} + \dfrac{1}{\alpha R_g}} i = R_g \left(\frac{\alpha}{1 + \alpha} \right) i,$$

or

$$i_g = \left(\frac{\alpha}{1 + \alpha} \right) i. \tag{54-1}$$

If, for instance, $\alpha = 1/9$, the current through the instrument is $1/10$ of the total and the effective sensitivity is reduced by a factor of ten. Similarly if $\alpha = 1/99$ the sensitivity is reduced by a factor of one hundred. A calibrated or direct reading current meter is usually equipped with a set of shunts to allow it to be used for different ranges of measurement.

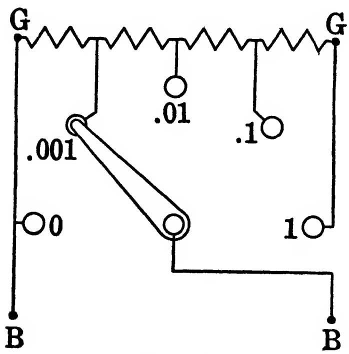


FIG. 90

In the adjustment of a bridge it is desirable to start with the galvanometer in a relatively insensitive condition and to increase its sensitivity as the condition of balance is approached. A convenient adjustable shunt in the form of a dial box is shown in Fig. 90. The terminals marked G are connected to the galvanometer, those marked B to the bridge. If the total resistance of the shunt is R_s and the resistance between the zero button and any other is βR_s, the galvanometer current for the 1-position of the contact arm is

$$i_g = \frac{R_s}{R_g + R_s} i,$$

and for any other position it is

$$i_g' = \frac{\beta R_s}{[R_g + (1 - \beta)R_s] + \beta R_s} i = \frac{\beta R_s}{R_g + R_s} i,$$

where i is the total current in either case, and R_g is the galvanometer resistance. The sensitivity for the second position relative to the first is $i_g'/i_g = \beta$. The factor β is usually taken as a tenth, a hundredth, and so on for successive positions, as shown in the figure. Since the relative sensitivity is independent of R_g this type of shunt is described as *universal*. In practice R_s should be large compared with R_g so that the sensitivity in position 1 is practically the full sensitivity of the unshunted galvanometer.

55. Resistance Bridge Measurements. — The most important applications of the bridge principle to resistance measurements are described below.

The Post Office Box. — This is a plug box with the resistance elements arranged as indicated in Fig. 91. Two keys (K) are mounted on the box, and four binding posts (P). Internal connections are shown by broken lines and external connections by

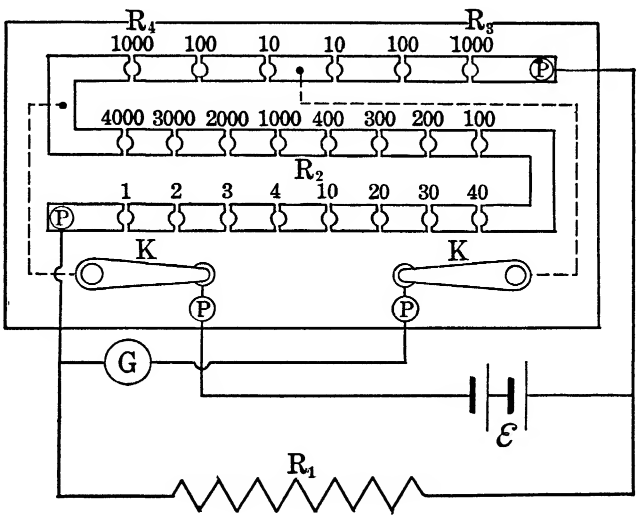


FIG. 91

full lines. The resistance to be measured is R_1. Comparing Fig. 91 with Fig. 86 we see that the circuit is a simple Wheatstone bridge. The keys are included to prevent continuous flow of current with the attendant heating. This is a necessary precaution in all bridge measurements, for steady heating even at a low rate will in time raise the temperature of the coils enough to change their resistances appreciably. In making an observation the battery key is closed *first* to permit a steady flow of currents to be established, then the galvanometer key is tapped. If there is a deflection the battery key is released and a readjustment of the bridge is made. This process is repeated until the gal-

vanometer shows no deflection when its key is opened and closed.

Returning to Fig. 91, the bridge arms indicated by R_3 and R_4 are known as the *ratio arms*, since they may be set in certain definite ratios, 100 : 1, 10 : 1, and so on down to 1 : 100. The remaining resistances R_2 constitute the adjustable arm. For any given R_1 there are usually several possible arrangements of the bridge. The primary consideration is to set the ratio arms so that R_2 is as large as possible, for the smaller the fraction of itself by which it can be varied the greater the accuracy of the balance. Any further choice in arrangement should be made in accordance with the principles for best sensitivity developed in article 53. The position of the galvanometer and the battery may be interchanged if necessary. For example, if R_1 is of the order of a thousand ohms we take $R_3 = 10$, $R_4 = 100$, or $R_3 = 100$, $R_4 = 1000$. In this case it makes little difference which pair of values we choose except that the latter pair allows the use of a larger battery. The galvanometer is as shown in the figure.

On the other hand if the resistance to be measured is of the order of ten ohms we set the ratio arms at 10 and 1000, and interchange the galvanometer and battery, so that the former is connected from the junction of the larger resistances to the junction of the smaller.

The Slide Wire Bridge. — This is a more accurate device than the post office box. The resistances R_3 and R_4 are replaced by a single straight resistance wire. The junction with the galvanometer arm is made through a sliding contact, which also serves as a tap key. Assuming that the resistance of the wire per unit length ρ is constant,

$$\frac{R_3}{R_4} = \frac{\rho l_1}{\rho l_2} = \frac{l_1}{l_2},$$

where l_1 and l_2 are the lengths into which the slider divides the wire. Therefore

$$R_1 = R_2 \frac{l_1}{l_2}, \tag{55-1}$$

and we require only one known resistance to determine R_1, the lengths l_1 and l_2 being obtained by direct measurement after the bridge is balanced. The construction of the slide wire unit is shown schematically in Fig. 92, together with the external connections. The wire W is connected to a heavy copper strip with

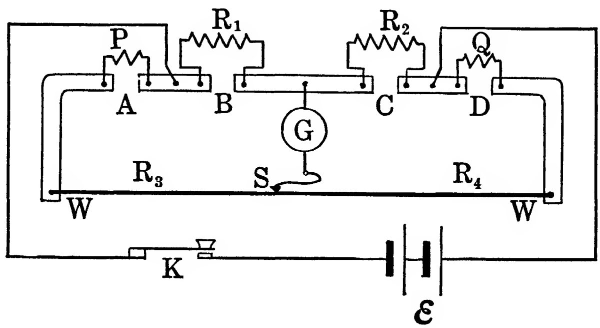


FIG. 92

four gaps. Into two of them, B and C, the unknown and the standard resistance are connected. The other gaps, A and D, may be short-circuited with metal straps, or may contain two auxiliary resistances P and Q whose purpose is explained below. The wire actually lies on a scale by means of which the ratio l_1/l_2 is determined. In making an observation the battery key K is closed first, then the key incorporated in the slider S.

The slide wire bridge may be balanced very accurately, so it is worth while to eliminate any possible errors associated with the wire itself. In the first place the ends of the wire may not coincide with the ends of the scale, or there may be a small resistance where the wire is connected to the copper strip. Let us suppose that the extra resistance at the zero end of the scale is equal to that of a section of the wire α units long. Similarly let β be the equivalent end correction at the other end of the scale. Then, short-circuiting P and Q and using known resistances for R_1 and R_2,

$$\frac{R_1}{R_2} = \frac{l_1 + \alpha}{l_2 + \beta}, \tag{55-2}$$

when the bridge is balanced. Next interchanging R_1 and R_2 and rebalancing,

$$\frac{R_2}{R_1} = \frac{l_1' + \alpha}{l_2' +} \tag{55-3}$$

From (55–2) and (55–3) α and β may be obtained by solution, since all other quantities are known. The ratio R_1/R_2 should be large, 10 : 1 or 20 : 1, in order to include as much of the wire as possible and thereby obtain the end corrections in terms of the mean resistance $\bar{\rho}$ per unit length; for the bridge wire is seldom absolutely uniform throughout its entire length. This is, in fact, the second source of error, to eliminate which we must calibrate the wire.

The best method of calibration is due to Carey-Foster. Let P and Q be two nearly equal resistances, their difference being some small fraction, a tenth or a twentieth, perhaps, of the resistance of the whole wire. For instance we may make P the effectively zero resistance of the metal strap, and Q a small resistance consisting of a short length of resistance wire soldered to two connecting lugs. R_1 and R_2 must be adjustable, but not necessarily known. Let us balance the bridge for any value of R_1/R_2; then interchange P and Q and rebalance. If the first balance point corresponds to a scale reading l_1 and the second to l_1', it is evident that the resistance of the portion of the wire in the interval $l_1 - l_1'$ equals $Q - P$, since there has been an exchange of equal amounts of resistance between the R_3 arm and the R_4 arm of the bridge. By manipulating R_1 and R_2 we may cause the interval $l_1 - l_1'$ to fall anywhere along the wire, thus dividing it into a number of sections of equal resistance. The lengths of the sections differ slightly due to lack of uniformity of the wire. By adding the lengths and dividing by the number of sections, the mean length is found, that is, the corresponding length of ideal uniform wire whose resistance per unit length has the mean value $\bar{\rho}$ of the actual wire. Subtracting the length of each section from the mean length and dividing by the length gives the correction per unit length over the given section. We

are now able to construct a calibration curve, plotting total correction against scale reading. By starting with a correction α for zero on the scale the end correction is included with the other, and the total length λ_1 of the mean wire equivalent to R_3 is obtained by adding to l_1 the corresponding correction as shown on the graph. If l is the scale length, the total equivalent length of $R_3 + R_4$ is given by

$$\lambda = l + \alpha + \beta,$$

and the equivalent length λ_2 for R_4 is then $\lambda - \lambda_1$. Therefore for an accurate determination of an unknown resistance (55–1) is replaced by $R_1 = R_2(\lambda_1/\lambda_2)$.

When the bridge is used to make a direct measurement of resistance as described above it is often called a *meter bridge*, because the slide wire is usually a meter long. Obviously, however, the bridge may be used to measure the difference between two nearly equal resistances by inverting the process used for the calibration of the wire. When thus used the bridge is called a *Carey-Foster bridge*. An important application is the determination of the temperature coefficient of resistance (See Table, p. 165). The method consists of measuring the difference between two approximately equal resistances, the temperature of one being kept constant and that of the other varied. From a series of observations the change of resistance per degree is obtained, and thence the temperature coefficient. As the resistances must not be removed from the temperature baths or other thermostats in which they are placed the transfer from one arm of the bridge to the other is accomplished by a mercury cup switching device of negligible resistance.

The Callendar and Griffiths Bridge. — This is a variety of slide wire bridge specially adapted to resistance thermometry, that is, the determination of temperature by measurement of the resistance of a coil at the given temperature. Pure platinum wire is used for the coil, which is enclosed in a long tube of porcelain or refractory metal capable of withstanding temperatures as high as 1200° C. By means of flexible leads the resistance element

T is connected into one arm of the bridge (Fig. 93). A pair of identical dummy or compensating leads is connected into the opposite arm to eliminate temperature and contact effects associated with the leads. In the same arm is a set of resistances arranged in the manner of a plug box. The other arms, R_1 and R_2, are equal. The slide wire and all other resistances, except T, are assembled as a unit in one case, the only external connections being galvanometer, battery and flexible leads.

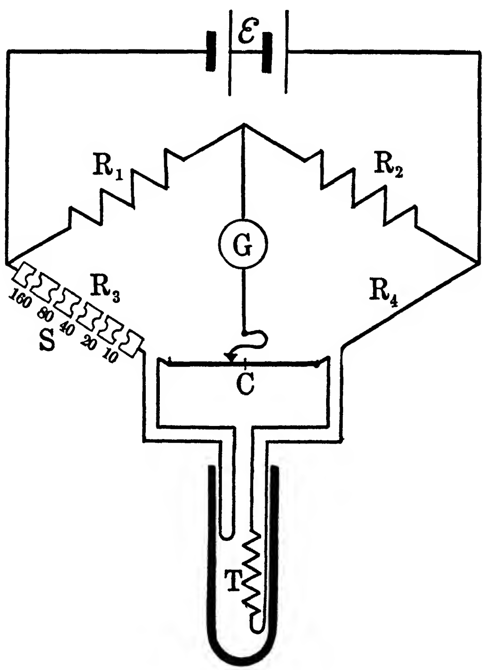


FIG. 93

It is customary in using this particular bridge to measure positions on the wire with reference to its electrical center C, a point that may be found by balancing the bridge with both sets of leads short-circuited and all adjustable resistance removed from R_3. If in making a measurement of T the balance point is a distance l from C,

where S is the amount of adjustable resistance included in R_3, r is the resistance of either set of leads, W is the total resistance of the wire, and ρ is its resistance per unit length. The above equation reduces to

$$T = S + \rho(2l). \tag{55-4}$$

As usually constructed the smallest resistance element in S has a resistance exactly equal to 20 cm of the wire. It is therefore equivalent to a 10 cm displacement of the slider, and is so marked. Similarly the second coil has a resistance equal to 40 cm of the wire, and is equivalent to a 20 cm displacement. If l' is the total equivalent displacement corresponding to S, (55-4) becomes simply $T = 2\rho(l + l')$. When not otherwise known the value of ρ may be found by some such method as that of Carey-Foster.

The actual calculation of the temperature may be performed in two ways. Callendar has shown that the resistance of a platinum coil at $t°$ C is accurately represented by

$$T = T_0(1 + \alpha t + \beta t^2), \tag{55-5}$$

where T_0 is the resistance at 0° C, and α and β are constants. These may be determined by measuring T at two known temperatures. The formula then gives implicitly at least the temperature corresponding to any measured value of the resistance. A simpler method is to neglect the βt^2 term in (55-5) and to define a *platinum temperature* t_p by

$$T = T_0(1 + \alpha' t_p),$$

or more explicitly,

$$t_p = 100 \frac{T - T_0}{T_{100} - T_0}, \tag{55-6}$$

the two scales being made to agree at 0° and 100°. To obtain the true temperature t an empirical formula is used, namely,

$$t - t_p = \delta\left\{\left(\frac{t}{100}\right)^2 - \left(\frac{t}{100}\right)\right\}.$$

It is usually convenient to plot the curve represented by this equation and find the correction for t_p graphically. The quan-

tity δ depends on the degree of purity of the platinum; for extremely pure samples it is very close to 1.5.

Galvanometer Resistance. — An ingenious application of the Wheatstone network is Kelvin's method of measuring galvanometer resistance (Fig. 94). The galvanometer is placed in one of the resistance arms, R_1 say, and serves as its own balance

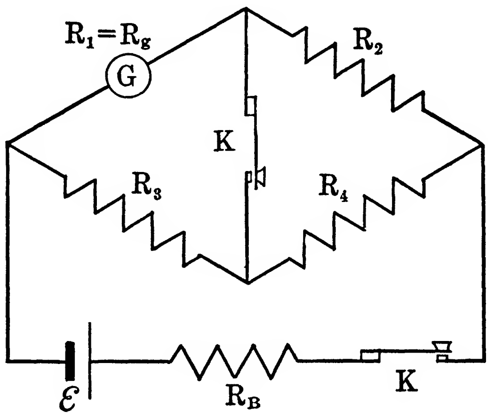


FIG. 94

indicator. In the regular galvanometer arm there is only a short-circuiting key. As opening or closing this key does not affect the bridge in any way when the latter is balanced, R_2, R_3 and R_4 are adjusted until the galvanometer deflection is the same with the key open as it is with it closed. Under this condition

$$R_g \qquad R_3$$

With a sensitive instrument a large resistance R_B must be placed in series with the battery to prevent excessive deflections.

The Kelvin Double Bridge. — Although the Wheatstone bridge has a wide range of utility, it is not well adapted to the measurement of very small resistances, that is, resistances of the order of one hundredth of an ohm or less. This is due in part to the fact that with the equipment usually available it is impossible to make a sensitive arrangement, but more particularly to the

difficulty of avoiding contact and junction resistances of the same order of magnitude as the resistance to be measured. The *double bridge* (Fig. 95) devised by Kelvin does not suffer from either of these defects. R_1 is the unknown resistance, and R_2 is a continuously variable resistance such as a slide wire. The connection between these, which has some small unknown resistance r, is shunted by a pair of resistances R_5 and R_6. It is this shunting of the junction resistance that distinguishes the double bridge from the simple Wheatstone bridge.

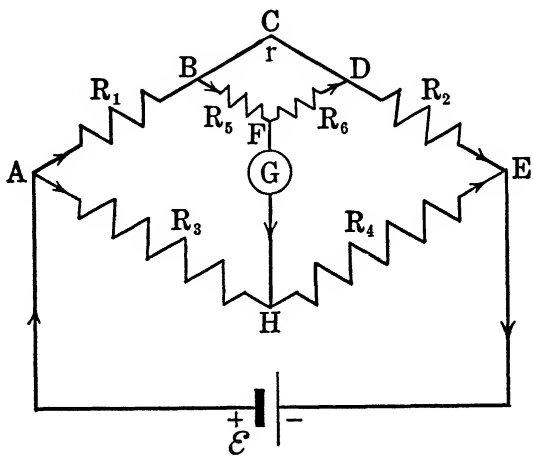


FIG. 95

To find the conditions for balance we observe that when there is no current through the galvanometer F and H are at the same potential. Therefore

$$R_1 i_1 + R_5 i_5 = R_3 i_3, \qquad R_6 i_6 + R_2 i_2 = R_4 i_4, \tag{55-7}$$

and

$$i_1 = i_2, \qquad i_3 = i_4, \qquad i_5 = i_6, \tag{55-8}$$

where i_1 is the current through R_1, and so on. Also, considering the parallel paths BCD and BFD,

$$i_5 = \frac{r}{r + R_5 + R_6} i_1, \qquad i_6 = \frac{r}{r + R_5 + R_6} i_2. \tag{55-9}$$

Combining (55–7) and (55–9) gives

$$\left(R_1 + \frac{rR_5}{r + R_5 + R_6}\right) i_1 = R_3 i_3,$$

$$\left(R_2 + \frac{rR_6}{r + R_5 + R_6}\right) i_2 = R_4 i_4.$$

Dividing one by the other leads with the aid of (55–8) to

$$\frac{R_1(r + R_5 + R_6) + rR_5}{R_2(r + R_5 + R_6) + rR_6} = \frac{R_3}{R_4},$$

or

$$(R_1R_4 - R_2R_3)(r + R_5 + R_6) + (R_4R_5 - R_3R_6)r = 0,$$

which is satisfied if

Evidently this set of conditions represents the only balance independent of the junction resistance r, an essential point.

The double bridge appears in several standard forms. A convenient one is shown schematically in Fig. 96. In order to

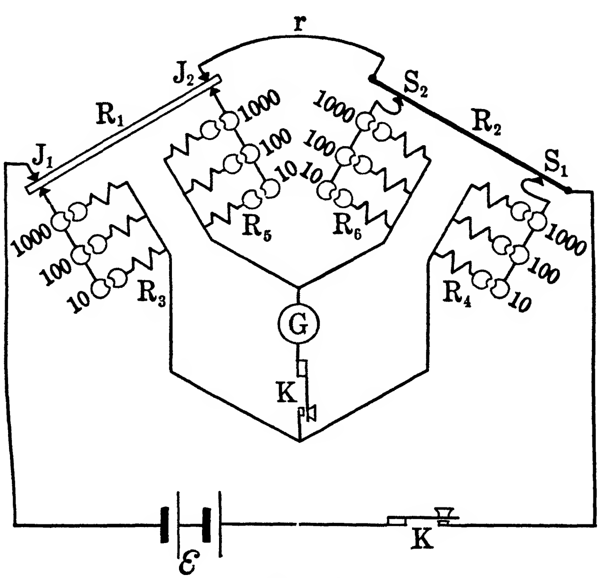


FIG. 96

avoid the double adjustment of ratios indicated by (55–10) the resistances R_3, R_4, R_5 and R_6 each consist of a set of coils with plugs, so arranged that the ratio $R_3/R_4 = R_5/R_6$ may be set at definite values 100 : 1, 10 : 1 and so on down to 1 : 100, as desired. The unknown resistance R_1, shown as a rod, is clamped in pairs of jaws, J_1 and J_2. The adjustable resistance R_2 consists of the portion of the slide wire included between two sliding contacts, S_1 and S_2. The end portions of the wire do not enter into the balance as they come into the junction resistance r and into the battery arm respectively.

Calibration of the wire is accomplished by dividing it into a number of sections of equal resistance, with the aid of a known resistance in place of R_1, in the manner described for the Carey-Foster bridge. A curve is plotted showing resistance per unit length, or total resistance from one end, along the wire, whereby the resistance in any interval S_1S_2 is readily determined. In making measurements of the unknown resistance, R_2 is taken as the mean of several observations in which S_1S_2 is allowed to fall on different parts of the wire. R_1 is then obtained on multiplying by the appropriate ratio factor.

An analysis of sensitivity similar to that carried out for the Wheatstone bridge shows that for best sensitivity a low resistance galvanometer should be used.

56. The Potentiometer. — Of equal importance to the measurement of resistance is the measurement of electromotive force or potential difference. For when $\mathcal{E}$ and R are known, the current i is determined and the electrical condition of a circuit is therefore completely specified. For accurate measurements it is essential that no current whatever be taken by the measuring apparatus, otherwise the e.m.f. measured may differ appreciably from the true value, which exists in the absence of the measuring equipment. This stringent requirement is met by a resistance network device called a *potentiometer*, whose main features are indicated in Fig. 97. A battery $\mathcal{E}_B$ causes a current i to flow through a resistance R_P. The magnitude of the current may

be varied by means of an auxiliary adjustable resistance R_B, in series with R_P. Across a certain portion r_1 of R_P is bridged a source of e.m.f. $\mathcal{E}_1$ and a galvanometer in series. Let us suppose that r_1 is chosen so that $i_g = 0$. Then evidently $r_1 i = \mathcal{E}_1$, for by Kirchhoff's second law $r_1(i + i_g) + R_g i_g = \mathcal{E}_1$, R_g being the gal-

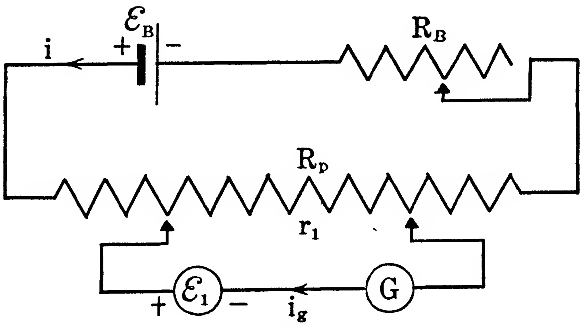


FIG. 97

vanometer resistance. Balancing the potentiometer consists therefore in adjusting r_1 until the galvanometer shows no deflection. Note that this cannot be done unless $\mathcal{E}_1$ is connected with the proper polarity. If $\mathcal{E}_1$ is replaced by another e.m.f. and a new balance obtained with r_2, *i being unchanged*, we see that

$$\frac{\mathcal{E}_1}{\mathcal{E}_2} = \frac{r_1}{r_2}. \tag{56–1}$$

That is, the potentiometer compares e.m.f.'s in terms of resistances. Supposing that $\mathcal{E}_1$ is due to a *standard cell* $\mathcal{E}_S$, and that $\mathcal{E}_2$ refers to any unknown e.m.f. $\mathcal{E}$, (56–1) becomes

$$\mathcal{E} = \mathcal{E}_S \frac{r}{r_S}. \tag{56–2}$$

The construction of standard cells of known e.m.f. is described in Chapter VI. The Weston standard cell, with an e.m.f. of 1.0183 volts * at 20° C, is the most common.

If the current through R_P is always kept the same we may

* This value is expressed in terms of the *international* volt, which differs slightly from the *absolute* volt. See Chapter XII.

mark the elements of which R_P is composed to read directly in volts. In several standard types of potentiometer therefore the standard cell is applied to a fixed portion of R_P only, and the balance effected by adjustment of R_B. This establishes a definite current in R_P and consequently a definite potential drop across each resistance element, so that when the unknown $\mathcal{E}$ is balanced by varying r the desired magnitude is indicated directly.

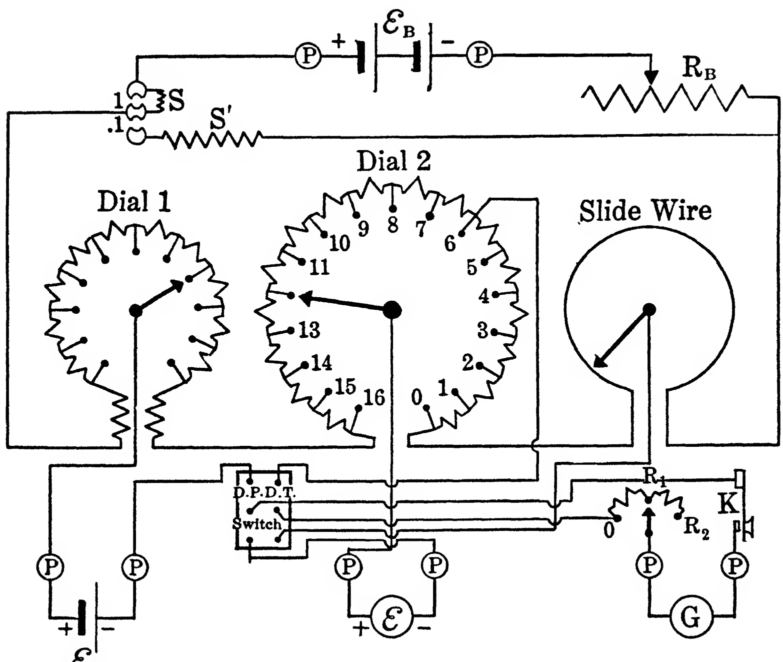


FIG. 98

The *Leeds and Northrup potentiometer* employs the above principle. The wiring diagram of this instrument is given in Fig. 98, the binding posts for external connections being indicated by *P*. The resistances in Dial 1, Dial 2, and the slide wire represent R_P. A double-pole double-throw switch is used to change from the standardizing position to the measuring. Thus when the switch is up in the figure $\mathcal{E}_S$ and G are across ten coils in Dial 2 and a certain number in Dial 1 depending on the value

of $\mathcal{E}_S$. The balance is obtained by varying R_B. The switch is now thrown down, placing $\mathcal{E}$ and G across part of Dial 2 and part of the slide wire, and balance is obtained by adjusting both of these. The slide wire is actually a long wire wound on a drum with a revolving contact and scale. The resistances S and S' are so arranged that with a plug in the normal position, marked 1, the current through R_P is 0.02 ampere and the figures

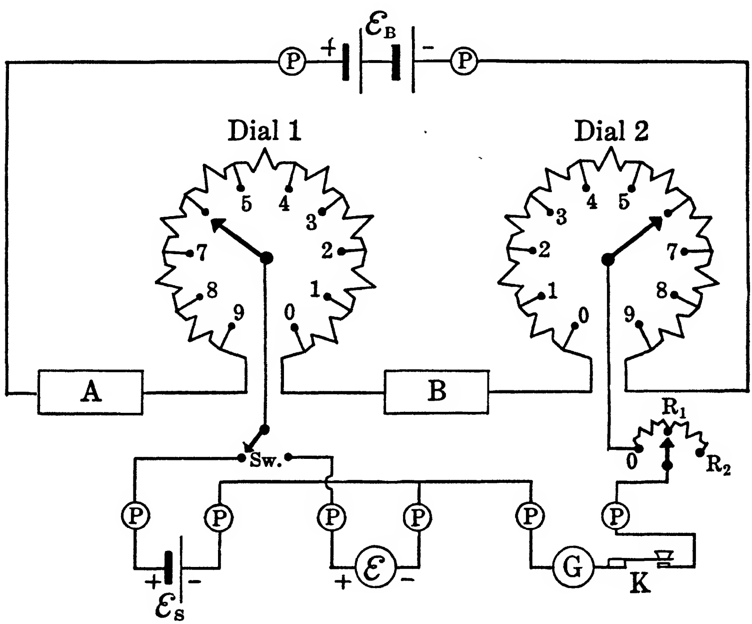


FIG. 99

on Dial 2 indicate tenths of a volt. Under these conditions the slide wire may be read directly to 0.00005 volt. When the plug is in the 0.1 position the current through R_P is reduced to 0.002 ampere and consequently the voltage indications are reduced to a tenth of their normal value. The high resistances R_1 and R_2 in series with G serve to protect the latter from excessive deflections and $\mathcal{E}_S$ from excessive current drain while adjustments for balance are being made.

A potentiometer which is not direct reading is somewhat

more flexible than the direct reading type and is therefore sometimes used. The *Wolff potentiometer* (Fig. 99) is a good example. Dial 1 has nine 1000 ohm resistances and Dial 2 nine 100 ohm resistances. The blocks A and B represent two identical plug box arrangements totaling 100 ohms each, with only *one* set of plugs. Plugs are moved only from one to the corresponding position in the other; so that the total included resistance of A plus B is always 100 ohms. The resistance of the entire potentiometer is always 10,000 ohms, and consequently the current is constant. Balancing consists of adjusting the dials and transferring plugs between A and B until G shows no deflection. The total bridged resistance, r_s or r as the case may be, is the sum of that shown on the dials and in B. Then $\mathcal{E}$ is calculated by (56–2).

As suggested at the beginning of the article, one of the most important uses of the potentiometer is the determination of current and the calibration of current indicators, the method being simply to introduce a known resistance into the given current circuit and measure $\mathcal{E}$ across it. There are also several other useful applications of the potentiometer which are described at appropriate places in later chapters.

CHAPTER VI

CHEMICAL AND THERMAL EFFECTS

57. **Electrolysis.** — Liquids, like solids, may be either insulators or conductors. Conducting liquids are generally called *electrolytes*, the most common examples being aqueous solutions of inorganic salts and acids. As the mechanism of electrolytic conduction is quite different from that of metallic conduction, we must investigate it in detail.

Consider first a solution of silver nitrate ($AgNO_3$) with two rods or plates of silver, called *electrodes*, partly immersed in it. Current is caused to pass through the electrolyte from one electrode to the other. The electrode by which the current enters the solution is designated as the *anode*, and the one by which it leaves as the *cathode*. We observe that, as the current flows, silver is deposited on the cathode, and incidentally an equal amount is removed from the anode, the electrolyte remaining unchanged. Furthermore, measurement shows that the mass of silver deposited is exactly proportional to the product of the current by the time for which it flows, that is, to the quantity of electricity passing through the electrolyte. The phenomenon just described is termed *electrolysis*. If now we repeat the experiment using a solution of copper sulphate ($CuSO_4$) and copper electrodes, we find that copper is deposited on the cathode and removed from the anode, the mass involved being as before proportional to the quantity of electricity which passes. However, the constants of proportionality, namely, the masses deposited per unit charge, are not the same in the two cases. They are, in fact, in the same ratio as the chemical equivalents of silver and copper. The *chemical equivalent* of a substance is the atomic weight, or in the case of a compound substance the group weight, divided by the valence.

Faraday, who was the first to make a systematic study of electrolysis, discovered that the simple relations described above represent general laws. Thus, we may write:

(I) *The mass of a substance liberated at an electrode is proportional to the total charge passing.*

(II) *The mass of a substance liberated at an electrode by a unit charge is proportional to the chemical equivalent of that substance.*

These laws may be expressed analytically in simple form. Let M be the mass of substance liberated, i the current, and t the time of flow. Then for the first law,

$$M = Zit, \tag{57-1}$$

where Z is a constant called the *electrochemical equivalent* of the given substance. Z is usually expressed in grams per coulomb. For the second law let w be the atomic, or group, weight of the substance involved, and v its valence. Then

$$Z = \frac{w}{v}\left(\frac{1}{F}\right). \tag{57-2}$$

The quantity $1/F$ is evidently the mass per coulomb for a substance of unit molecular weight and valence. As the atomic weight of hydrogen is 1.008 and its valence is unity, F, called *Faraday's constant*, is the number of coulombs required to liberate 1.008 grams of hydrogen. Its numerical value is 96,490. Observe that 96,490 coulombs will liberate a mass in grams of any substance numerically equal to its chemical equivalent. This amount of a substance is called a *gram-equivalent.* Evidently F is measured in coulombs per gram-equivalent.

Although Faraday discovered the laws of electrolysis, the explanation of the underlying physical mechanism is due to Arrhenius. According to his theory of *electrolytic dissociation* a certain number of the molecules of the dissolved substance are separated into two or more parts, each of which bears a charge. These charged carriers are called *ions.* Under the influence of an applied e.m.f. they move to the electrodes, the positively charged *cations* to the cathode and the negatively charged *anions*

to the anode. At the electrodes the ions give up their charge and the substance of which they are composed is deposited. The continuous arrival of positive charge at one electrode and negative charge at the other is equivalent to a transfer of charge from one to the other, and constitutes the current passing through the electrolyte. Material deposited may actually appear on an electrode, or it may be involved in some secondary chemical reaction. For example, the ions in the silver nitrate solution are Ag^+ and NO_3^-. The Ag^+ is deposited as metallic silver on the cathode, thereby increasing the latter's mass. The NO_3^- is deposited on the anode, but it cannot exist alone in the neutral state; so it combines with the silver of which the anode is composed. The silver nitrate thus formed goes back into solution and the net result is loss of silver by the anode. The working of Faraday's laws of electrolysis is often confused by these secondary reactions, particularly in cases where the material of the electrodes is different from that of the ions. Thus if current passes through a solution of sulphuric acid (H_2SO_4) in water, the electrodes being of platinum, hydrogen is deposited on the cathode and appears directly in small bubbles. The anion SO_4^{--} is deposited on the anode, but cannot exist alone in the neutral state, nor can it combine with the platinum. It therefore removes H_2 from a water molecule to form sulphuric acid again, releasing oxygen which appears on the anode in bubbles. This is an interesting case because the water is continually decomposed, the concentration of the solution becoming therefore steadily greater. We shall return to this case later.

The question now arises as to the magnitude of the charge carried by the individual ions. We know that a gram-equivalent of any substance carries 96,490 coulombs. Moreover, the number of atoms or molecules in a gram-equivalent depends only on the valence. For a monovalent substance the number is $6.06(10)^{23}$, the number of atoms in 1.008 grams of hydrogen. For a bivalent substance the number is one half of this, and so on. These facts indicate that every monovalent ion carries the same basic charge, positive or negative as the case may be. A bivalent

ion carries two of these charges, a trivalent ion three. We suspect at once, of course, that the magnitude of this basic charge is equal to that of the electron and the proton, $1.59(10)^{-19}$ coulomb. Dividing 96,490 by $6.06(10)^{23}$ does indeed give that exact value, which we may regard as a confirmation not only of the theory of electrolysis but also to some extent of the entire electron theory of matter.

Problem 57a. Given that the valence of silver is one and that the valence of copper is two, calculate the chemical equivalents of these metals and also of the *acid radicals* NO_3 and SO_4. Ans. 107.88, 31.79; 62.01, 48.03.

Problem 57b. Calculate the electrochemical equivalents of Ag and Cu. Ans. 0.001118o gm/coulomb, 0.0003294 gm/coulomb.

Problem 57c. Calculate the electrochemical equivalent of oxygen and determine for how long 1 ampere must flow through a dilute solution of sulphuric acid to liberate 1 gram of this gas. How much hydrogen is produced at the same time? Ans. 3 hr 21 min, 0.126 gm.

58. Conductivity of Electrolytes. — The electrical conductivity of an electrolyte must be proportional to the number of ions per unit volume, just as the conductivity of a solid is proportional to the number of free electrons. If the number of molecules of dissolved substance per unit volume of solution is c, the *concentration*, and the fraction of these molecules ionized is δ, the *degree of dissociation*, the conductivity σ is given by $\sigma = Cc\delta$, where C is a constant of proportionality. In electrolysis it is customary to define the *equivalent conductivity* λ as the ratio of the true conductivity to the concentration. Thus,

$$\lambda \equiv \frac{\sigma}{c} = C\delta. \tag{58–1}$$

This quantity is independent of the concentration if δ is constant, but experiment shows that in general λ increases with decreasing c, approaching a limiting value asymptotically as c converges to zero. We suppose that this limiting value λ_0 corresponds to complete dissociation, that is, $\delta = 1$. Then $\lambda_0 = C$, and

$$\frac{\lambda}{\lambda_0} = \delta. \tag{58–2}$$

As σ, and hence λ, may be determined experimentally as a function of c, and the results extrapolated to give λ_0, (58–2) leads to an experimental value of δ. This is a matter of some interest as δ may be found in other ways, which do not involve electrical measurements or theories. Values of δ found by means of (58–2) agree well with the results of the other methods, an added confirmation of Arrhenius' theory. The variation of δ with c for a silver nitrate solution at 18° C is shown in the adjoining table.

$c(10)^{-17}$	δ
0	1.000
12	0.970
61	0.934
121	0.910
242	0.875
485	0.833
606	0.817

The measurement of the conductivity of an electrolyte is quite simple if the passage of current does not cause any change in the composition of the electrolyte or of the surface of the electrodes. In such a case the measurement may be made by any ordinary direct current method. Usually, however, the measurement of conductivity is complicated by a phenomenon known as *polarization*. Consider for example sulphuric acid in water, with platinum electrodes. As we have seen (art. 57), hydrogen is liberated at the cathode and oxygen at the anode. A layer of gas collects on each electrode, so that in effect the electrodes are converted from platinum to hydrogen and oxygen respectively. But dissimilar electrodes in a solution with which they can react constitute a voltaic cell (art. 51), so there is an internal e.m.f. of magnitude $\mathcal{E}_p$, which opposes the external or applied e.m.f. Unless the external e.m.f. is greater than $\mathcal{E}_p$ current cannot be made to pass through the cell. As the work performed in the decomposition of a given amount of water must be equal to the energy obtained on recombination, $2.873(10)^5$ joules per gram-molecule, we can easily calculate a minimum possible value for $\mathcal{E}_p$. Suppose a gram-molecule of water is decomposed. This requires the passage of (2×96490) coulombs through the cell, and corresponds to electrical work of amount $(2 \times 96490)\mathcal{E}_p$. This latter quantity is not equal to the energy of formation,

however, for experiment shows that the cell absorbs heat as the decomposition takes place, $0.797(10)^5$ joules per gram-molecule at 20° C. Thus if no energy is dissipated,

$$(2 \times 96490)\mathcal{E}_p + 0.797(10)^5 = 2.873(10)^5,$$

which gives $\mathcal{E}_p = 1.08$ volt. Actually it takes about 1.7 volts to produce steady decomposition of water in the given cell. The polarization voltage is therefore greater than is necessary to supply energy for the decomposition. The excess voltage, called the *over-voltage* of the cell, is ascribed to a dissipation of energy attendant on the formation and release of bubbles of gas at the electrodes.

When polarization is present measurement of conductivity is usually made by a method due to Kohlrausch. The electrodes are contained in two flasks which are connected by a straight glass tube (Fig. 100). The area of the electrodes is made large

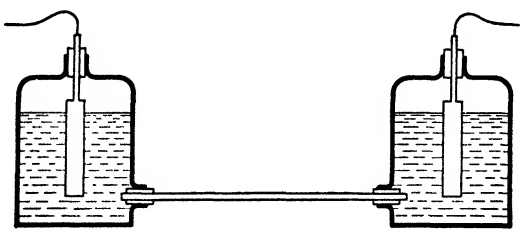


FIG. 100

in order to increase the time necessary for $\mathcal{E}_p$ to build up. The entire unit is inserted in one arm of a slide wire bridge, whose applied e.m.f. is not steady, but alternating. That is, it reverses direction periodically. The current through the electrolytic cell also reverses direction, and if the frequency of reversal is great enough — several hundred alternations per second — polarization is unable to develop to any appreciable extent. The measured resistance is then the true resistance of the electrolytic cell. Of course, a special alternating current galvanometer, or a telephone receiver, must be used in balancing the bridge, instead

of the ordinary direct current instrument described in Chapter V. In order to eliminate the unknown resistance of the electrodes and the electrolyte in the flasks, measurement is first made with a long tube between the flasks, then with a short tube of the same cross-section. The difference between the two resistance values obtained is the resistance of a column of liquid of length equal to the difference of the tubes. Knowing the cross-section of the tubes the conductivity of the electrolyte may be calculated at once.

Electrolytic conductivity like metallic conductivity is independent of e.m.f. and current, that is, electrolytes obey Ohm's law. The dependence on temperature, however, is different in the two cases; for metallic conductivity decreases slowly as the temperature rises, while electrolytic conductivity increases quite rapidly, as much as two or three per cent per degree centigrade at room temperature.

In order to apply Ohm's law to a cell which is subject to polarization we must write

where $\mathcal{E}$ is the applied e.m.f. and r_0 is the true ohmic resistance. If, now, we define the *apparent resistance* r by the equation $i = \mathcal{E}/r$ we see that

$$r = r_0 + \frac{\mathcal{E}_p}{i}. \qquad (58\text{-}3)$$

As $\mathcal{E}_p$ depends in a complex manner not only on the current but also on the length of time it has flowed and on the previous condition of the electrodes, r is in general variable and to some extent indeterminate. This result evidently applied to a voltaic cell, itself causing the flow of current, as well as to a simple electrolytic cell with an external applied e.m.f., which explains the statement made about r in article 51.

59. The Voltameter. — Electrolytic cells are frequently used for the absolute measurement of current. It is only necessary to determine the mass M of substance liberated at an electrode in a time t. Then by the first law of electrolysis the current is

$$i = \frac{M}{Zt}, \qquad (59\text{–}1)$$

where Z is the electrochemical equivalent of the substance liberated. We usually choose a cell in which a metal is deposited directly on the cathode without any secondary chemical reaction. A cell used for current measurements is called a *voltameter*.

For ordinary laboratory use a copper voltameter is convenient and sufficiently accurate. The electrolyte is a solution of copper sulphate made by dissolving copper sulphate crystals in about four times their weight of water. The electrodes are copper plates, usually arranged as shown in Fig. 101 in order to utilize both sides of the cathode. It is necessary to determine the increase in mass of the cathode rather than the decrease in mass of the anode, as the latter may contain impurities which dissolve out during electrolysis.

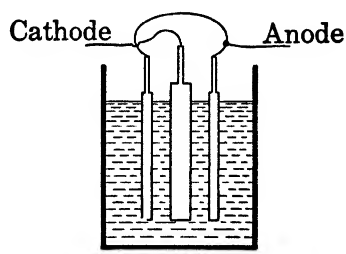


Fig. 101

The cathode, having been cleaned and weighed, is placed in the cell and current is allowed to pass for a known time. In order to obtain a hard smooth deposit the current density at the cathode should not exceed 20 milliamperes per square centimeter. The cathode is now removed, washed in distilled water and weighed again. As the electrochemical equivalent of copper is 0.0003294 gm/coulomb we have all the necessary data for the calculation of i by (59–1).

For the most accurate absolute determinations of current a silver voltameter is invariably used. The electrolyte is a solution of silver nitrate, about 15 or 20 parts by weight of silver nitrate

to 100 parts of water. The anode is, of course, of silver. The cathode, however, is usually a platinum cup which serves as a container for the electrolyte (Fig. 102). In order to prevent impurities from reaching the cathode a small porcelain cup *P* — porous to the electrolyte — surrounds the anode. The voltameter shown schematically in Fig. 102 is the Richards' type. There are several other types which differ considerably in details.

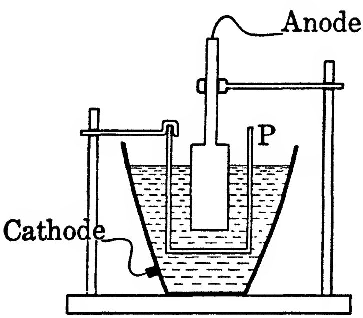


Fig. 102

To obtain the highest accuracy, a few parts in 100,000, great care in the preparation and assembly of the apparatus is essential. For a detailed discussion of the technique of voltameter measurements the reader should consult some reference work. (See for example, Glazebrook: *Dictionary of Applied Physics*, Vol. II, p. 247 ff.) Because of the high accuracy obtainable under proper conditions the silver voltameter is used to establish the primary standard current at the Bureau of Standards and at other standardizing laboratories. That is, the ampere is defined in terms of the amount of silver deposited by it per second. This quantity, the electrochemical equivalent of silver, is 0.001118o gm/coulomb.

60. Theory of the Voltaic Cell. — The physical characteristics of voltaic cells have been described in article 51. As these cells are simply electrolytic cells in which the electrodes react with the electrolyte in such a way as to produce an internal e.m.f., their electrochemical behavior is governed by Faraday's laws of electrolysis. It remains, however, to investigate the mechanism by which the internal e.m.f. is produced. This was long a much debated mystery, whose eventual solution is credited to Nernst.

In order to understand Nernst's theory it is necessary first to familiarize ourselves with the phenomenon of *osmosis*. Suppose

we have a U-tube (Fig. 103) with water in one arm and an aqueous solution of some sort in the other. The liquids are separated by a semipermeable membrane or plug M which permits free passage of water but not of the dissolved substance in the solution. It is found that the height of the liquid is not the same in the two arms of the tube, being greater for the solution than for the pure water. The pressure p_s just to the right of M is greater than the pressure p_w to the left. Now p_s is made up of two partial pressures, due to the water and to the dissolved substance respectively. Moreover, as water passes through M freely, its partial pressure on the right of M must be equal to the total pressure p_w on the left; so that $p_s = p_w + p$, where p is due to the dissolved substance. The quantity p, which is thus the difference of pressure across M, is called the *osmotic pressure*. A few measurements of p under different conditions suffice to show that in a dilute solution at least it is proportional to the absolute temperature and to the number of dissolved molecules or ions per unit volume of solution. In other words the dissolved particles behave as if they constituted an ideal gas, quite independent of the solvent. With this fact we may return to Nernst's theory.

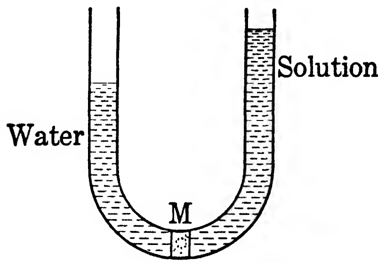


FIG. 103

When a metal electrode is dipped into an electrolyte there is a tendency for positively charged metallic ions to go into solution. According to Nernst we can measure this tendency in terms of a pressure P called the *electrolytic solution pressure*. If some of the metallic ions are present in the solution before the electrode is introduced we are able to distinguish three cases. Thus if P is greater than p, the osmotic pressure of the ions in the solution, ions must pass into solution from the electrode, leaving the latter negatively charged. If P is equal to p nothing occurs,

while if it is less ions are deposited on the electrode, charging it positively. If the electrolyte does not contain ions of the same substance as the electrode, p is zero, of course. In any case equilibrium occurs when the electric forces between the electrode and the ions are just sufficient to balance the difference between P and p. For example, a silver electrode dipped in a solution of silver nitrate becomes positive relative to the solution, for the electrolytic solution pressure of silver is very small. The difference of potential between an electrode and a solution can be measured by suitable means, and serves to confirm Nernst's theory.

We are now in a position to understand the difference between a simple electrolytic cell, without resultant e.m.f., and a voltaic cell, which has an e.m.f. If two identical electrodes are placed in an electrolyte each in general differs in potential from the electrolyte, but by the same amount, so there is no difference of potential between the electrodes. On the other hand, electrodes of different materials differ in potential from the solution by different amounts, and therefore differ in potential from each other, this potential difference being the e.m.f. of the cell. Basically, therefore, the electrolytic cell is a special case of the voltaic cell rather than the other way, as we have heretofore considered it.

Incidentally it appears from the solution pressure theory that we can produce an e.m.f. without chemical reactions. The potential difference between an electrode and an electrolyte containing the same substance depends on the osmotic pressure of the metallic ion, which in turn depends on the concentration. If, then, a cell with identical electrodes is arranged so that the concentration of the solution differs as we pass from one electrode to the other an e.m.f. is produced and we have a *concentration cell.* The e.m.f. of a concentration cell is usually quite small. Thus a silver nitrate cell with silver electrodes has an e.m.f. of about 0.05 volt at ordinary temperatures if the concentration at one electrode is ten times as great as it is at the other.

There are several ways in which cells may be classified. For

theoretical purposes it is usually best to classify them as *reversible* and *irreversible*. A reversible cell is one which may be restored to its original condition after discharge by passing a reverse current through it, the energy required for charging being exactly equal to that obtained from the cell during discharge. An irreversible cell does not have this property, either because energy is dissipated or because there are secondary chemical actions which cannot be reversed. For practical purposes it is convenient to classify cells as *primary* and *secondary*. A primary cell is used for discharge only. It is usually, but not necessarily, irreversible. A secondary cell, often called a storage cell, is alternately charged and discharged, and is necessarily reversible, therefore. Various cells will be described under this classification in following articles.

There is an important theorem relative to reversible cells, originally deduced by Gibbs and by Helmholtz independently. Let us suppose that a given reversible cell at an absolute temperature T generates Q units of electricity. If H is the *heat of formation* of the chemical compounds which appear in the cell during the process, that is, the energy released during the formation of the compounds, and H' is the external heat absorbed, the electromotive force $\mathcal{E}$ is determined by the equation,

$$Q\mathcal{E} = H + H'. \qquad (60\text{–}1)$$

This follows immediately from the law of conservation of energy.

Let us now apply a cyclic process to the cell in order to establish a relation between heat absorbed and work performed. Thus suppose the cell first generates Q units of charge at temperature T absorbing external heat H'. This is represented graphically by ab in Fig. 104. Now insulate the cell thermally and pass a reverse current through it. If heat was absorbed before, it must now be evolved. It cannot escape due to the thermal insulation, so the temperature of the cell rises. We limit this step in the cycle, represented by bc, to infinitesimal changes. Thus the temperature rises to $T + dT$ and the e.m.f.

changes to $\mathcal{E} + d\mathcal{E}$. The insulation is now removed and the reverse current continued until we reach the point d, $H' + dH'$ units of heat being given up at the temperature $T + dT$. The point d is chosen in such a way that if we replace the thermal insulation and allow the cell to generate charge the cell is restored

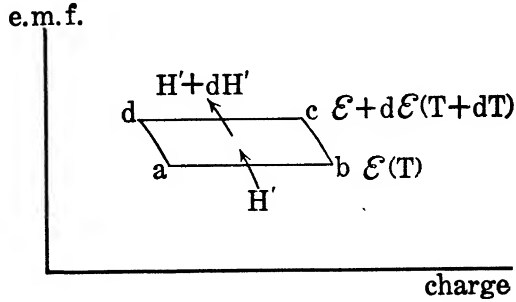


FIG. 104

exactly to its original condition when the temperature reaches the value T. Since we have a complete cycle the energy dH' given up by the cell in the form of heat must be equal to the energy received electrically. To calculate this electrical energy we observe that the total charge produced by the cell during the cycle equals the total charge absorbed by it and therefore, as bc and da are infinitesimals, ab equals cd. Thus the electrical energy received is $Q(\mathcal{E} + d\mathcal{E}) - Q\mathcal{E} = Qd\mathcal{E}$ and we have

$$dH' = Qd\mathcal{E}. \tag{60–2}$$

The cycle is reversible since we are dealing with a reversible cell, so we may give the cell dH' units of heat and receive $Qd\mathcal{E}$ units of electrical work if we choose.

This reversibility leads us to a very interesting fact. Suppose we have two reversible cells working in cycles with the same temperature limits. We may arrange the cells so that, in effect, one cell runs the other backwards, that is, the external work $Q_1d\mathcal{E}_1$ supplied by one is exactly equal to the amount $Q_2d\mathcal{E}_2$ absorbed by the other. Thus $dH_1' = dH_2'$. Also $H_1' = H_2'$; otherwise we can run the combination of cells in a direction such

that heat is removed from bodies at a temperature T and is delivered to other bodies at a temperature $T + dT$ without expenditure of energy, a procedure contrary to one of the basic laws of thermodynamics. Hence $dH_1'/H_1' = dH_2'/H_2'$, and as we are dealing with *any* two reversible cells it is evident that the ratio dH'/H' is the same for *all* reversible cells, for all reversible mechanisms of any sort, in fact, working between the temperatures T and $T + dT$. As dH'/H' is independent of material properties it offers an ideal means of establishing an absolute temperature scale, and, as it happens, the *thermodynamic* or *Kelvin scale* to which we are accustomed is obtained by making T proportional to H'. That is,

$$\frac{dT}{T} = \frac{dH'}{H'}. \qquad (60\text{–}3)$$

Combining this with (60–2) gives

$$H' = QT\frac{d\mathcal{E}}{dT}, \qquad (60\text{–}4)$$

and finally, using (60–1),

$$Q\mathcal{E} = H + QT\frac{d\mathcal{E}}{dT},$$

or

$$\mathcal{E} = h + T\frac{d\mathcal{E}}{dT}, \qquad (60\text{–}5)$$

where $h \equiv H/Q$ is the heat of formation per unit charge.

The *Gibbs-Helmholtz* equation (60–5) shows us that the electrical energy developed by a cell at constant temperature is equal in amount to the energy supplied by the internal chemical processes only if $\dfrac{d\mathcal{E}}{dT} = 0$. Otherwise the cell absorbs heat as it generates charge, or evolves it, depending on whether $\dfrac{d\mathcal{E}}{dT}$ is positive or negative.

61. Primary Cells. — Several different types of primary cell have been devised to meet various practical requirements. The

most important feature in any of them is the depolarizing mechanism, for the primary chemical process almost always involves polarization effects which must be neutralized by secondary reactions of some sort. A cell that polarizes appreciably is of little practical value.

The Daniell Cell. — In its simplest form this cell has a zinc electrode in a solution of sulphuric acid and a copper electrode in a solution of copper sulphate, the two liquids being separated by a porous cup (Fig. 105). The zinc is amalgamated to eliminate the effect of impurities which cause a local electrolytic action that wears away the electrode. When the cell is in action SO_4^{--} ions combine with zinc at the electrode to form $ZnSO_4$ while the H^+ ions move out through the walls of the cup. On reaching the copper sulphate they displace copper ions and form H_2SO_4 with the remaining SO_4^{--}, the copper being deposited on the copper electrode. The copper sulphate solution is thus the depolarizer, for it prevents hydrogen, the most common source of polarization, from appearing at the electrode.

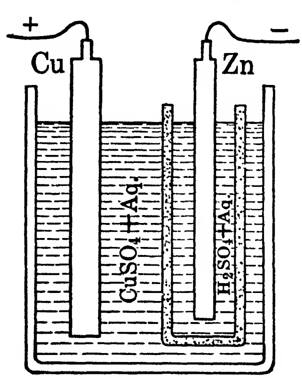


FIG. 105

Use of the porous cup may be avoided by placing the zinc electrode above the copper, and relying on the different densities of the solutions to keep them from mixing. A Daniell cell so arranged is often called a *gravity cell.* In any form of Daniell cell mixture will take place by diffusion if the cell is allowed to stand idle. This cell is therefore adapted only to a service where there is a continuous flow of current. The internal resistance of the cell is rather large, of the order of an ohm.

Ideally, at least, the Daniell cell is reversible, so we may calculate its e.m.f. by means of the Gibbs-Helmholtz equation. As $\frac{d\mathcal{E}}{dT}$ is very close to zero (60–5) reduces to $\mathcal{E} = h$. The resultant action in the cell is the formation of zinc sulphate and

the decomposition of copper sulphate, so *h* is the heat of formation of $ZnSO_4$ minus that of $CuSO_4$. The first of these heats in joules per coulomb is 0.82 and the second − 0.27, so that $\mathcal{E} = 1.09$ volt. This is the observed value, subject to slight variation with different concentrations.

The Leclanché Cell. — In its basic form this cell has zinc and carbon electrodes in a solution of ammonium chloride (NH_4Cl). The carbon electrode is in a porous cup packed with manganese peroxide (MnO_2) which is the depolarizer (Fig. 106). When current flows Cl^- ions move to the zinc electrode where they form zinc chloride. The NH_4^+ ions move to the carbon where they separate into ammonia (NH_3) and hydrogen. The latter reacts with the MnO_2, forming Mn_2O_3 and water. As the depolarizing action is rather slow the cell becomes somewhat polarized if used continually, but recovers on being allowed to stand idle for a time. It is, therefore, particularly suited to intermittent service. Its e.m.f. is 1.5 volts.

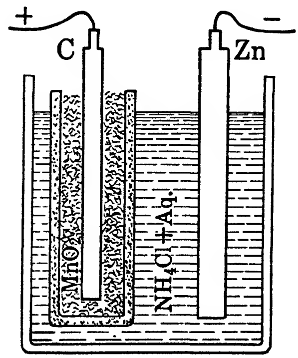


Fig. 106

The common *dry cell* is a special form of Leclanché cell. In this case both the depolarizer and the electrolyte are made up into a paste. The cell is contained in a zinc can which serves as one electrode, the whole being sealed to prevent the paste from drying out. The internal resistance of a new dry cell is one-tenth of an ohm or less, so fairly large currents may be drawn momentarily before polarization develops.

The Weston Standard Cell. — When carefully prepared and cared for this cell provides a highly accurate standard of electromotive force. The somewhat complicated construction of the cell is indicated in Fig. 107. The cell is contained in a sealed glass tube of H-form. Platinum wires sealed into the glass make contact with the electrodes, which are mercury and cadmium

amalgam respectively. Above the mercury is a paste composed of mixed cadmium sulphate and mercurous sulphate, which is the depolarizer. The electrolyte is a solution of cadmium sulphate, crystals being included to insure saturation. The e.m.f. at 20° C is 1.0183 volts * with a negative temperature coefficient of 4 parts in 100,000.

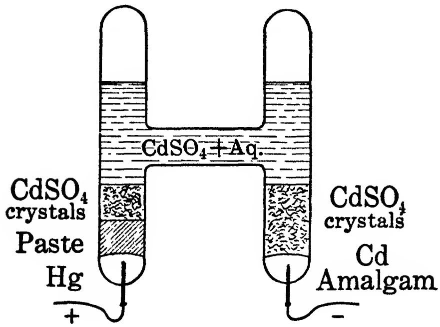


FIG. 107

Extensive research has been devoted to the preparation of the Weston cell, a full discussion of which may be found in Glazebrook: *Dictionary of Applied Physics*, Vol. II, p. 260 ff. In use, care should be taken to maintain the entire cell at a uniform temperature, and to protect it from agitation. Measurements are always made by a potentiometer method (art. 56) to avoid appreciable flow of current.

62. Secondary Cells. — There are two types of secondary cell of practical value.

The Lead Cell. — Of the two types this is distinctly the more important. It was devised by Planté in 1859 and has been highly developed commercially since that time. When the cell is fully charged the electrodes are lead oxide (PbO_2) and lead in a spongy form, respectively. The electrolyte is a solution of sulphuric acid with a specific gravity of 1.20 to 1.28. Usually, in order to obtain a large capacity, each electrode consists of a number of *plates* connected together (Fig. 108). Wooden *separators* prevent plates of opposite polarity from coming into contact.

As the cell discharges SO_4^{--} ions move to the negative (lead) plate with which they combine to form $PbSO_4$. The action at the positive plate is more complex. The H^+ ions reduce the PbO_2 to PbO, forming water in the process. Then the PbO

* International volts. See Chapter XII.

combines with the sulphuric acid forming $PbSO_4$ again, and more water. As lead sulphate appears on both plates according to the foregoing explanation of the chemical reactions, this explanation is called the *double sulphate theory*. Some objections have been raised against it but the weight of evidence is in its favor. The reactions in charging are of course the exact reverse of those described above.

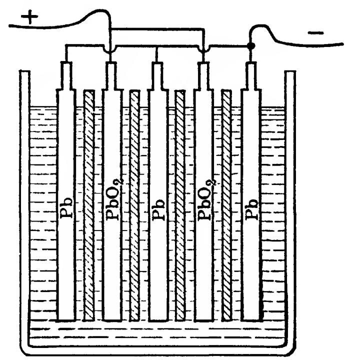


FIG. 108

Since acid is removed from the electrolyte during discharge, the specific gravity of the solution falls. This provides a convenient indication of the state of charge, as the specific gravity is easily measured with a *hydrometer*. When this quantity has fallen to about 1.15 the cell should be considered fully discharged, for at lower values the lead sulphate tends to change from the normal form in which it is deposited to an insoluble form which is not decomposed on charging. This not only destroys part of the active material but, by covering the plates, causes a serious increase in the internal resistance of the cell. A cell which has suffered in this way is said to be *sulphated*. Prolonged idleness also causes sulphation unless the cell is fully charged.

The e.m.f. at full charge is 2.05 volts or slightly more. The discharge value is about 1.80 volts. The internal resistance is 0.01 ohm or less if the cell is not sulphated, and there is no appreciable polarization, so very large currents may be obtained. The capacity of a cell is usually expressed in terms of the quantity of electricity it can generate, the unit being the *ampere-hour*. That is, a 100 ampere-hour cell will provide 10 amperes for 10 hours, or 5 amperes for 20 hours, and so on. The ordinary cell varies in capacity from 50 to 200 ampere-hours. For such a cell a current which charges or discharges it in eight or nine hours is considered to have a good working value. As has been pointed

out a cell should not be discharged beyond a certain point, but it may be overcharged without any particular harm except the decomposition of water in the electrolyte. This loss may be repaired by adding the proper amount of *distilled* water.

The construction of the plates is a matter of some interest. Originally Planté *formed* the plates electrolytically by placing pure lead in dilute sulphuric acid and passing current first in one direction and then in the other. This process is slow and expensive, so that other methods have been devised. The simplest method consists of casting skeleton plates, or *grids*, of lead and filling these with a paste which is easily converted by electrolysis to the desired form. For the positive plates the paste may be made of Pb_3O_4 and for the negative plates it may be made of PbO. Pasted negatives are quite satisfactory and are widely used. Pasted positives, although frequently used, are somewhat less desirable as they have a tendency to crack or buckle.

For both commercial and laboratory use several cells — usually three — are connected in series and assembled in one case, the resulting unit being called a *storage battery*, or an *accumulator.*

The Iron-Nickel Cell. — This cell has electrodes of iron and of nickel oxide. The electrolyte is a solution of potassium hydroxide. It is not described in detail as it is not extensively used. It has however certain valuable characteristics. For a given energy content it is lighter than a lead cell. Moreover it is capable of withstanding considerable abuse both electrically and mechanically. It is particularly adapted therefore to electric vehicles and to special uses such as military field service. When charged the cell has an e.m.f. of about 1.4 volts.

Problem 62a. The temperature coefficient of the lead storage cell is 0.0004 volt per degree centigrade. How much heat at 20° C does the cell absorb per ampere-hour of discharge? Ans. 100 cal.

63. Thermo-Electricity. — It is possible to produce an e.m.f. by thermal means without resort to chemical reactions. Seebeck in 1826 discovered that a current flows in a circuit composed of

two metallic conductors of different materials if the temperatures of the two junctions are different. There is therefore an e.m.f. in such a circuit, dependent on thermal conditions, called a *thermal electromotive force*, or a *thermo-electric force*. The combination of conductors in which this e.m.f. is generated is known as a *thermocouple*. The energy expended as the current flows is supplied by the absorption of heat from external sources, assuming the temperature at every point in the circuit is maintained unchanged, for there is no other possible means of supply. The first experimental observation of this phenomenon was made by Peltier, who discovered that when current flows across a junction between two metallic conductors at constant temperature heat is absorbed or evolved, depending on the direction of the current. The quantity of the *Peltier heat* is proportional to the quantity of charge which crosses the junction. It depends also on the temperature of the junction and the materials of the two conductors. Thus when current flows in a thermocouple circuit due to a temperature difference of the junctions, heat is absorbed at the high temperature junction and evolved, in a smaller amount, at the low temperature junction. The difference between the two quantities is converted into electrical energy by the thermocouple. The Peltier heat is perfectly reversible. If a reverse current is forced through the thermocouple by means of an external e.m.f., heat is absorbed at the low temperature junction and evolved at the high. Care must be taken to distinguish between the reversible Peltier heat and the irreversible *Joule heat* due to resistance. We need not include the latter in the thermodynamical discussion, as we can make it negligibly small by using small currents and conductors of large cross-section. This follows from the fact that the Joule heat is generated at a rate equal to Ri^2, R being the resistance of the circuit and i the current, while the Peltier heat is developed at a rate proportional to i alone.

The existence of the Peltier heat at the junction of two conductors indicates the presence there of an e.m.f. Suppose a charge Q passes across the junction in the positive direction,

that is, the direction in which the e.m.f. tends to produce current flow. If Π is the e.m.f., the electrical work generated is $Q\Pi$, and this must be equal to the heat H_p absorbed. Thus Π equals the Peltier heat per unit charge. We can understand the production of the e.m.f. at the junction, in a general way at least, from our knowledge of the structure of conductors (art. 49). In the interior of a metal there are free electrons which, taken collectively, behave like a gas. When two metals are placed in contact the electron gas in each diffuses into the other, but the densities differ in general, so the rates of diffusion are different, and the metals become oppositely charged. As the charges are built up an opposing electric field develops and equilibrium is soon established. The differential diffusion thus creates an e.m.f., which is measured internally from the negative metal to the positive, as that is the direction in which it tends to produce current in a closed circuit.

The discussion of the junction e.m.f. contains a valuable suggestion. If the temperature is different at two points in a conductor the density of the electron gas must differ. We may therefore expect an e.m.f. in a conductor whenever there is a temperature gradient. This effect, which is easily found experimentally, was first predicted by Thomson (Lord Kelvin) and is called the *Thomson effect*. As the electron density is greater at lower temperatures the Thomson e.m.f. should be from cold to hot. In some metals however it is in the other direction, an anomaly which has not been satisfactorily explained. When current flows there is a reversible absorption or evolution of heat just as in the case of the junction e.m.f. The Thomson e.m.f. between two points on a conductor which differ in temperature by dT may be expressed in the form σdT, where σ, called the *specific heat of electricity*, is some function of the temperature.

We are now able to express the e.m.f. of a thermocouple analytically, and to deduce some important theorems. Denoting the conductors of which the thermocouple is composed by a and b, the total e.m.f. is $\mathcal{E}_{ab}$, measured around the circuit in such a

direction that a positive current flows from a to b at the warm junction. If the upper junction in Fig. 109 is the warm one the positive direction of e.m.f., and current, is indicated by the arrow. Let the temperature of the warm junction be T_1 and that of the cool junction be T_2. Then

$$\mathcal{E}_{ab} = [\Pi_{ab}]_{T_1} - \int_{T_2}^{T_1} \sigma_b dT - [\Pi_{ab}]_{T_2} + \int_{T_2}^{T_1} \sigma_a dT$$

$$= \left|\Pi_{ab}\right|_{T_2}^{T_1} + \int_{T_2}^{T_1} (\sigma_a - \sigma_b) dT. \quad (63\text{–}1)$$

The first important property of $\mathcal{E}_{ab}$ is apparent if we choose some temperature T intermediate between T_1 and T_2 and calculate the e.m.f.'s $\mathcal{E}_{ab}'$ and $\mathcal{E}_{ab}''$ for the temperature intervals $T_1 - T$ and $T - T_2$ respectively. We have

$$\mathcal{E}_{ab}' = \left|\Pi_{ab}\right|_{T}^{T_1} + \int_{T}^{T_1} (\sigma_a - \sigma_b) dT,$$

$$\mathcal{E}_{ab}'' = \left|\Pi_{ab}\right|_{T_2}^{T} + \int_{T_2}^{T} (\sigma_a - \sigma_b) dT.$$

Adding these,

$$\mathcal{E}_{ab}' + \mathcal{E}_{ab}'' = \left|\Pi_{ab}\right|_{T_2}^{T_1} + \int_{T_2}^{T_1} (\sigma_a - \sigma_b) dT = \mathcal{E}_{ab}. \quad (63\text{–}2)$$

We may evidently divide $T_1 - T_2$ into any number of intervals with the same result. Hence, *the e.m.f. of a thermocouple for*

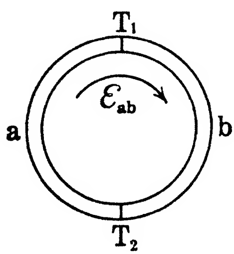


FIG. 109

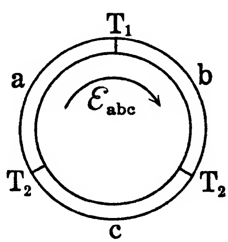


FIG. 110

any temperature interval $T_1 - T_2$ is the sum of the e.m.f.'s corresponding to any smaller intervals into which $T_1 - T_2$ may be subdivided.

Next let us introduce a third metal c, and compare the thermocouples ac, bc and ab. For the temperature interval $T_1 - T_2$

$$\mathcal{E}_{ac} = |\Pi_{ac}|_{T_2}^{T_1} + \int_{T_2}^{T_1} (\sigma_a - \sigma_c)dT,$$

$$\mathcal{E}_{bc} = |\Pi_{bc}|_{T_2}^{T_1} + \int_{T_2}^{T_1} (\sigma_b - \sigma_c)dT,$$

and by subtraction,

$$\mathcal{E}_{ac} - \mathcal{E}_{bc} = -\,|\Pi_{bc} + \Pi_{ca}|_{T_2}^{T_1} + \int_{T_2}^{T_1} (\sigma_a - \sigma_b)dT, \quad (63\text{–}3)$$

as $\Pi_{ac} = -\Pi_{ca}$.

Now considei a closed circuit abc composed of the three metals, all parts of which are kept at the same temperature. There can be no resultant e.m.f. in this circuit; otherwise current would flow and heat energy would be converted directly into electrical or mechanical work without any temperature difference available, which is contrary to the laws of thermodynamics. As there are no temperature gradients the Thomson e.m.f.'s are all zero and we must have

$$\Pi_{ab} + \Pi_{bc} + \Pi_{ca} = 0. \quad (63\text{–}4)$$

Combining this with (63–3) gives

$$\mathcal{E}_{ac} - \mathcal{E}_{bc} = |\Pi_{ab}|_{T_2}^{T_1} + \int_{T_2}^{T_1} (\sigma_a - \sigma_b)dT = \mathcal{E}_{ab}. \quad (63\text{–}5)$$

Therefore, *if we know the thermal e.m.f. for each of two metals in combination with a third, we may obtain the e.m.f. for the combination of the two metals by subtraction.* This is very convenient, for it allows us to measure and tabulate e.m.f.'s for all metals against some one chosen as a standard. The e.m.f. for any pair of metals is then given by the difference of the corresponding tabular values. Lead is usually chosen for the reference metal as its Thomson coefficient is zero.

Lastly, let us include a third metal c in the thermocouple circuit ab (Fig. 110), keeping the junctions bc and ca at the

same temperature T_2, however. The e.m.f. is

$$\mathcal{E}_{abc} = \left[\Pi_{ab}\right]_{T_1} - \int_{T_2}^{T_1} \sigma_b dT + \left[\Pi_{bc}\right]_{T_2} + \int_{T_2}^{T_2} \sigma_c dT + \left[\Pi_{ca}\right]_{T_2} + \int_{T_2}^{T_1} \sigma_a dT$$

$$= \left[\Pi_{ab}\right]_{T_1} + \left[\Pi_{bc} + \Pi_{ca}\right]_{T_2} + \int_{T_2}^{T_1} (\sigma_a - \sigma_b) dT.$$

Using (63–4) we obtain

$$\mathcal{E}_{abc} = \left|\Pi_{ab}\right|_{T_2}^{T_1} + \int_{T_2}^{T_1} (\sigma_a - \sigma_b) dT = \mathcal{E}_{ab}. \qquad (63\text{–}6)$$

This result may evidently be extended to include any number of added metals. Thus, *any number of intermediate metals may be included in a thermocouple circuit without affecting its e.m.f., provided each such conductor has the same temperature at both ends.* This is the most important property of thermocouple circuits from a practical standpoint; for it permits the introduction of instruments, copper connecting wires and other accessories which must be included in the circuit for measurement purposes, without any disturbing effect.

It remains to determine the form of $\mathcal{E}_{ab}$ as a function of the absolute temperature T. Suppose the thermocouple works between the temperatures T and $T + dT$. As the Peltier and Thomson heats are both reversible we have a device whose thermodynamical properties are identical with those of the reversible voltaic cell. Therefore the fraction of the heat passing through the thermocouple which is converted into electrical energy must be equal to dT/T, as explained in article 60. This means

$$\frac{d\Pi_{ab} - \sigma_b dT + \sigma_a dT}{\Pi_{ab}} = \frac{dT}{T},$$

or

$$\frac{d\Pi_{ab}}{dT} - \frac{\Pi_{ab}}{T} + \sigma_a - \sigma_b = 0,$$

which may be expressed in the convenient form

$$\frac{d}{dT}\left(\frac{\Pi_{ab}}{T}\right) + \frac{\sigma_a - \sigma_b}{T} = 0. \qquad (63\text{–}7)$$

This equation establishes a general relation between the Peltier and Thomson coefficients. Now experiment shows that σ is very nearly proportional to T for most substances. Let us set

$$\left.\begin{aligned} \sigma_a &= k_a T, \qquad \sigma_b = k_b T, \\ \sigma_a - \sigma_b &= k_{ab} T, \end{aligned}\right\} \tag{63-8}$$

where $k_{ab} \equiv k_a - k_b$. This enables us to integrate (63–7) obtaining

$$\Pi_{ab} = k_{ab} T(T_{ab} - T). \tag{63-9}$$

The quantity T_{ab}, which is equal to the constant of integration divided by k_{ab}, is called the *neutral temperature*.

Finally, using (63–1), (63–8) and (63–9) we have for the total e.m.f.

$$\begin{aligned} \mathcal{E}_{ab} &= \left| k_{ab} T(T_{ab} - T) \right|_{T_2}^{T_1} + \int_{T_2}^{T_1} k_{ab} T dT \\ &= k_{ab} T_{ab}(T_1 - T_2) - k_{ab}(T_1^2 - T_2^2) + \tfrac{1}{2} k_{ab}(T_1^2 - T_2^2) \\ &= k_{ab}(T_1 - T_2)[T_{ab} - \tfrac{1}{2}(T_1 + T_2)]. \end{aligned}$$

Since in practice we always keep one junction at a fixed temperature, we may set $T_2 = T_r$, a constant reference temperature, and $T_1 = T$, a variable quantity. Then

$$\mathcal{E}_{ab} = k_{ab}(T - T_r)[T_{ab} - \tfrac{1}{2}(T + T_r)]. \tag{63-10}$$

It appears that $\mathcal{E}_{ab}$ is a parabolic function of the temperature (Fig. 111), being zero when $T = T_r$, of course, and again when $T = 2T_{ab} - T_r$, that is, when the *average* temperature of the junctions $\frac{1}{2}(T + T_r)$ equals T_{ab}. Curves are shown in the figure for various values of T_r. Evidently $\mathcal{E}_{ab}$ has a maximum for any value of T_r when $T = T_{ab}$.

As (63–10) depends on (63–8) it is sometimes not sufficiently accurate when the temperature varies over a wide range. In such a case the thermocouple must be calibrated. On the other hand, when small temperature differences only are involved, as is often the case, (63–10) serves very well. We are then interested in the rate of change of $\mathcal{E}_{ab}$. This quantity, called the

thermo-electric power, is given by

$$\frac{d\mathcal{E}_{ab}}{dT} = k_{ab}(T_{ab} - T). \tag{63–11}$$

Since T_r does not appear the fixed temperature is of no consequence. This is indicated graphically in Fig. 111 by the parallel tangents, whose slope is the thermo-electric power. In order that $\frac{d\mathcal{E}_{ab}}{dT}$ shall be large it is necessary that T be kept well away from T_{ab}. Equation (63–11), being linear, offers the most convenient means of recording data on thermocouples, either graphically

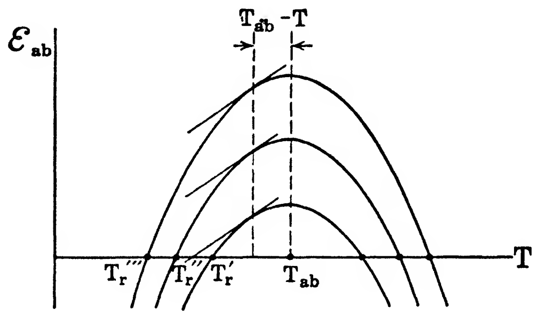


FIG. 111

or in tabular form. The thermo-electric power for various metals in combination with lead is determined experimentally. Then the value for any pair of metals is obtained by subtraction, as may be seen by differentiating (63–5). For tabulation it is convenient to use the centigrade scale. Denoting centigrade temperature by t, $T = t + 273°$ and (63–11) may be written

$$\frac{d\mathcal{E}_{ml}}{dt} = k_{ml}(t_{ml} - t) \equiv \alpha + \beta t, \tag{63–12}$$

where m indicates any metal, l indicates lead, and $-\alpha/\beta \equiv t_{ml}$. As thermal e.m.f.'s are usually very small, $\mathcal{E}_{ml}$ is measured in microvolts. Values of α and β for some metals and a few alloys are given in the adjoining table, which is based on the very complete work of Bridgman.* The values of the constants vary

* American Acad. of Arts and Sciences, Vol. 53, p. 269, (1918).

considerably with the condition of the substances, and must be regarded as approximate only. Note that according to our definition positive e.m.f. causes current to flow from m to l at the *warm* junction. Practice varies in this regard.

The thermal electromotive force for any metal in combination with lead expressed in terms of α and β is obtained by integrating (63–12). This gives

$$\mathcal{E}_{ml} = \alpha(t - t_r) + \tfrac{1}{2}\beta(t^2 - t_r^2), \qquad (63\text{–}13)$$

where t_r is the fixed reference temperature of one junction.

Substance	α	β
Bismuth	74.42	− 0.0320
Cadmium	− 12.002	− 0.3238
Cobalt	17.32	0.0780
Copper	− 2 777	− 0.00966
Iron	− 16.18	0.0178
Magnesium	0.095	− 0.00008
Nickel	17.61	0.0356
Platinum	3.092	0.02668
Silver	− 2.566	− 0.00864
Tin	− 0.230	0.00134
Tungsten	− 1.594	− 0.03410
Zinc	− 3.047	0.00990
Constantan	34.76	0.0794
Manganin	− 1.366	− 0.000828

Problem 63a. Show that

$$\Pi_{ab} = T\frac{d\mathcal{E}_{ab}}{dT},$$

$$\sigma_a - \sigma_b = -T\frac{d^2\mathcal{E}_{ab}}{dT^2},$$

are general relations applying to all thermocouples. Verify these relations for the special case in which $\mathcal{E}_{ab}$ has the form given by (63–10).

Problem 63b. Referring to the table of thermo-electric powers, calculate Π_{ab} and $\sigma_a - \sigma_b$ for a bismuth-cadmium thermocouple at 0° C. Ans. 0.024 joule/coulomb, $-79.7(10)^{-6}$ volt/degree.

Problem 63c. Express the thermal e.m.f. of a copper-constantan couple in terms of centigrade temperatures, the cool junction being at 0° C. Ans. $\mathcal{E}_{ab} = -37.54t - 0.0445t^2$ (microvolts).

64. Thermocouple Measurements and Instruments. — Thermocouples are of great value for the measurement of temperatures and temperature intervals of any magnitude up to 1600° C. The type of thermocouple is determined by the use for which it is intended. Up to 300° C copper and constantan — a copper-nickel alloy — are satisfactory. Iron and constantan will serve up to 800° C. For the higher temperatures platinum and some platinum alloy must be used. Temperature measurements in commercial practice usually consist in obtaining the thermal e.m.f., either directly by means of a potentiometer or from the deflection of a high resistance galvanometer, and referring the result to a calibration curve. In the laboratory it is often necessary to determine small changes in temperature accurately. In this case a potentiometer is always used to measure the e.m.f.'s to avoid a possible error due to the Peltier heat at the junctions of the thermocouple. One junction is kept at any convenient fixed temperature while the other varies. The change in e.m.f. divided by the thermo-electric power gives the corresponding temperature change.

Thermocouples are used, indirectly, to measure several physical quantities in addition to temperature. For example, if radiant energy in the form of heat or light falls on one junction it produces a rise of temperature and consequently an e.m.f., which may be used to measure the intensity of the radiation. Great sensitivity may be obtained if a number of thermocouples are connected in series (Fig. 112), alternate junctions being exposed to the heating influence so that the e.m.f.'s add. A device of this sort is called a *thermopile*.

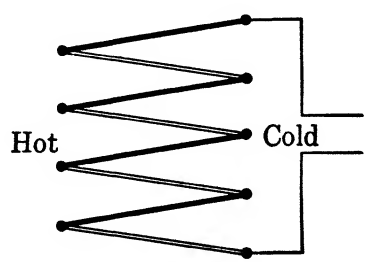


FIG. 112

In alternating current electrical measurements thermocouples play an important role, especially for high frequencies where the coil type of measuring instrument is not effective. A

thermocouple meter consists of a sensitive direct current meter element, a thermocouple and a small resistance unit, all included in a single case. The external current to be measured is passed through the resistance or *heater* unit. This is so arranged that the heat generated warms one junction of the thermocouple. The resulting thermal e.m.f. produces a direct current through the meter element. In most direct current indicators the deflection is proportional to the current, which in this case is proportional to the thermal e.m.f. This latter quantity is proportional to the average rate at which heat is generated, that is, to the mean value of the square of the alternating current. Thus the scale of a thermocouple meter is not uniform but is more extended at the upper end than at the lower. A convenient arrangement for a thermocouple meter is shown in Fig. 113. The meter element is designated by *M*. The thermocouple and heater are formed jointly by crossing two conductors, *ac* and *bd*, such as iron and constantan for example, and welding or soldering them at the point of contact *o*. The external current flows along *aob*, which serves as the heater. The thermocouple proper is *cod*.

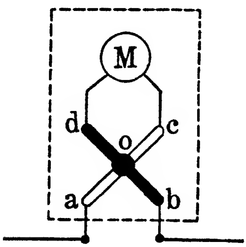


FIG. 113

65. Contact Potential Difference.—Although we have studied various phenomena dependent on the existence of free electrons in the interior of a conductor, we have as yet made no mention of the restraining forces which must exist at the surface of the conductor to oppose the escape of these electrons. There are several effects which depend on these surface forces.

In the first place, experiment shows that a definite amount of energy ϕ must be expended in carrying an electron from the interior of a conductor out through the surface. The quantity ϕ, called the *work function*, depends primarily on the substance of which the conductor is composed. It varies slightly with the temperature.

Now suppose that we place two conductors *a* and *b* (Fig. 114) in contact. If we carry a unit positive charge around a closed path such as that indicated in the figure, the total *electrical* work performed is zero since the forces are of an electrostatic nature

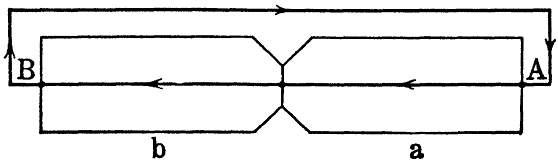


FIG. 114

(art. 48). Let V_{AB} be the difference of potential, if any, existing between a point A just outside the surface of conductor *a* and a point B just outside the surface of conductor *b*. The work equation is

$$-V_{AB} + \frac{\phi_a}{e} - \Pi_{ab} - \frac{\phi_b}{e} = 0,$$

or

$$V_{AB} = \frac{\phi_a - \phi_b}{e} - \Pi_{ab},$$

where Π_{ab} is the Peltier coefficient and e is the charge of an electron (a negative quantity). V_{AB} is known as the *contact potential difference* of the two conductors. Its existence is readily confirmed by experiment, although accurate measurement is very difficult. Its magnitude is of the order of a few tenths of a volt, so Π_{ab} is entirely negligible and we may write

$$V_{AB} = \frac{\phi_a - \phi_b}{e}. \tag{65–1}$$

Care must be taken to distinguish between the potential difference just described and that existing between an electrode and an electrolyte, to which the same name is sometimes applied.

66. Thermionic Emission. — The work function ϕ, described in the preceding article, is the controlling factor in an effect of great practical importance called *thermionic emission.* When a

conductor is heated electrons are emitted, the number increasing with great rapidity as the temperature rises. This phenomenon is utilized by many devices which involve electron currents, such as X-ray tubes, radio vacuum tubes, alternating current rectifiers, and electron relays. The mechanism of the emission is easily understood when we consider the properties of the electron gas in the interior of the conductor. The free electrons move about with a random distribution of velocities, their mean kinetic energy, however, being always proportional to the absolute temperature. Whenever an electron with kinetic energy greater than ϕ happens to move toward the surface from a point just inside, it is able to escape, losing an amount ϕ of kinetic energy in the process. The higher the temperature, the greater is the number of electrons with the necessary energy for escape, so the emission increases as the temperature rises.

From the kinetic theory of gases, and from thermodynamics, it is possible to calculate the rate at which electrons are emitted as a function of the temperature. This calculation was first made by Richardson, who showed that if ϕ is a constant the emission current per unit area must have the form

$$j = A'T^{1/2}e^{-\frac{b'}{T}}, \tag{66–1}$$

where T is the absolute temperature and A', b' are constants for any substance, b' being proportional to ϕ. If variation of ϕ with temperature is taken into account the emission current equation is

$$j = AT^2e^{-\frac{b}{T}}. \tag{66–2}$$

It is difficult to distinguish between the two equations experimentally with the range of temperatures available in the laboratory. However, (66–2) is superior theoretically and is therefore to be preferred. It is interesting that A is apparently a universal constant for pure metals, being equal to 60.2 when j is measured in amperes per square centimeter.

Electron emission usually takes place in a vacuum, the

emitting body being in the form of a *filament* F (Fig. 115) heated by a current from an external battery $\mathcal{E}_F$. Included in the vacuum space is another conductor P, usually called the *plate*. When this is maintained at a positive potential V relative to F by a plate battery $\mathcal{E}_P$, the emitted electrons are drawn away from the filament as they appear and a steady emission current exists. Evidently if the plate is at a negative potential no current can pass. In order to obtain the full emission current predicted by (66–2) it is necessary that V be great enough to remove the electrons from the vicinity of F very rapidly, otherwise a *space charge*

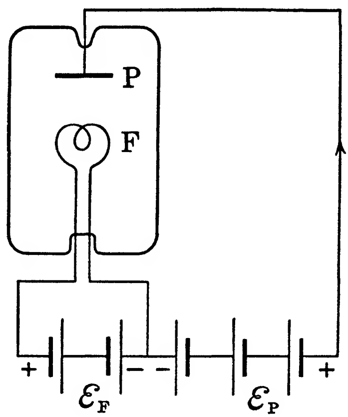


Fig. 115

(art. 80) accumulates between P and F and limits the current to the plate. This effect is shown in Fig. 116 where the full line curve, representing (66–2), gives the current i to the plate when V is very large. The broken curves show the currents actually obtained for smaller values of V. Evidently the higher the temperature, the greater must be the value of V to obtain the full emission current.

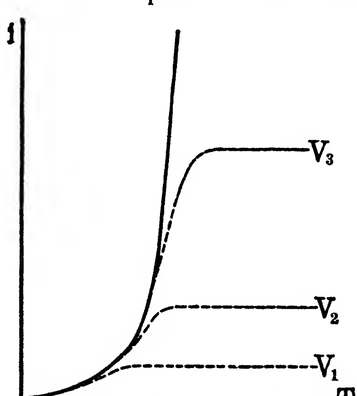


Fig. 116

The magnitude of the emission from any substance is greatly affected by the condition of the surface. For example, the presence of a minute amount of thorium on the surface of tungsten increases the emission many thousand times. Use is made of this fact in constructing filaments. Platinum filaments coated with various barium and strontium salts also have very high

emission. Values of A and b for a few substances are given in the table.

Substance	$A\left(\frac{\text{amp}}{\text{cm}^2}\right)$	b (deg. abs.)
Carbon	60.2	46500
Molybdenum	60.2	51500
Platinum	60.2	59000
Tungsten	60.2	52400
Thoriated Tungsten	3.0	30500

Problem 66a. For pure tungsten filaments the normal working temperature is about 2400° K (absolute). Calculate the emission current at this temperature. Ans. 0.115 amp/cm^2.

Problem 66b. The normal temperature for thoriated tungsten filaments is about 2000° K. Calculate the emission current. Compare it with the emission current for pure tungsten at the same temperature. Ans. 2.86 amp/cm^2, 2840 times as great.

67. Piezo-Electricity and Pyro-Electricity. — When an anisotropic crystal, such as quartz, is subjected to a mechanical stress it usually shows an electric polarization which manifests itself in the appearance of bound charges, positive on some faces of the crystal and negative on others. These charges are referred to as *piezo-electricity.* If an electric field is applied to an unstressed crystal, producing an electric polarization, the corresponding mechanical strain appears, this effect being the converse of the one first described. Piezo-electricity has found a number of applications, particularly in connection with the conversion of mechanical vibrations into electrical, or *vice versa.*

When a crystal with piezo-electric properties is warmed or cooled it also develops bound charges, which are designated by the term *pyro-electricity.* Presumably these charges are really piezo-electric, the strains which produce them being caused by the thermal expansion of the crystal. As regards the explanation of piezo-electricity itself, it is probable that the crystal is always polarized, but that due to the very slight conductivity surface

charges distribute themselves so as to neutralize the bound charges. When the crystal is deformed the condition of polarization is slightly changed and the surface charges no longer produce complete neutralization. If the strain is maintained for a sufficient time the surface charges can readjust themselves and the piezo-electricity should disappear, an effect which is observed experimentally.

CHAPTER VII

MAGNETIC FIELD OF A CURRENT

68. Scalar and Vector Products. — So far our use of vectors has been confined to the addition and subtraction of directed quantities. In the study of the magnetic fields due to currents and of induced electromotive forces certain products of vectors play such an important part that the treatment of the subject becomes very clumsy if we limit ourselves to the scalar methods employed for the most part in the preceding chapters.

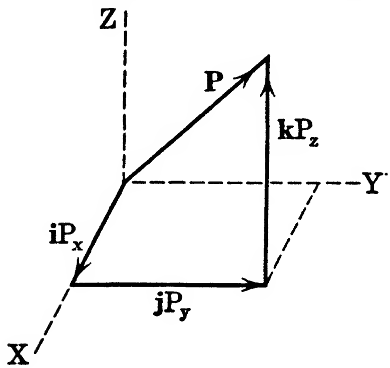


FIG. 117

In order to represent a vector in terms of its rectangular components we employ unit vectors i, j, k directed along the X, Y, Z axes respectively. If, then, the magnitudes of the rectangular components of a vector **P** are designated by P_x, P_y, P_z we have

$$\mathbf{P} = iP_x + jP_y + kP_z, \tag{68-1}$$

as is evident at once from Fig. 117.

Writing a second vector **Q** in the same form,

$$\mathbf{Q} = iQ_x + jQ_y + kQ_z, \tag{68-2}$$

and the sum of the two vectors is given by

$$\mathbf{P} + \mathbf{Q} = i(P_x + Q_x) + j(P_y + Q_y) + k(P_z + Q_z), \tag{68-3}$$

since each component of the sum is equal to the sum of the corresponding components of the two vectors.

Two products of vectors are important in physics, the scalar product and the vector product.

Scalar Product. — The scalar product of two vectors **P** and **Q** is defined as the product of the magnitude of **P** by that of **Q** by the cosine of the angle between them. It is a scalar quantity and is designated by $\mathbf{P} \cdot \mathbf{Q}$, the dot representing the cosine of the angle between the positive directions of the two vectors. It is equal to the component of either vector in the direction of the other multiplied by the magnitude of the other.

Many examples of the scalar product occur in physics. If a body on which a force **F** is acting suffers a displacement $d\mathbf{l}$ in a direction making an angle θ with the force the work done is

$$dW = F \cos \theta dl = \mathbf{F} \cdot d\mathbf{l}. \tag{68-4}$$

A plane surface may be represented as a vector directed along its positive normal and having a magnitude equal to its area. Therefore the electric flux through an element of surface $d\mathbf{s}$ (the element ds may always be considered plane even if the surface of which it forms a part is curved) is

$$dN = \mathbf{E} \cdot d\mathbf{s},$$

and Gauss' law (8–5) may be written

$$\int_s \mathbf{E} \cdot d\mathbf{s} = 4\pi \int_\tau \rho d\tau. \tag{68-5}$$

Next consider a current, whether it be due to a flow of electricity or of a material fluid. The current density **j** is a vector having the direction of the current and a magnitude equal to the current per unit cross-section. In terms of **j** the current di through an element of surface $d\mathbf{s}$ is $\mathbf{j} \cdot d\mathbf{s}$. For the current through ds (Fig. 118) is that through the cross-section $dA = ds \cos \theta$. Consequently,

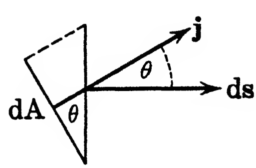


Fig. 118

$$di = jdA = j \cos \theta ds = \mathbf{j} \cdot d\mathbf{s}.$$

Hence the current through any surface s is given by the integral

$$i = \int_s \mathbf{j} \cdot d\mathbf{s}. \tag{68-6}$$

Vector Product. — This product is of great importance in electro-magnetism. The vector product of two vectors **P** and **Q** is a vector normal to their plane in the sense of advance of a right-handed screw rotated from the first to the second of the two vectors through the smaller angle between them. Its magnitude is equal to the product of the magnitudes of the two vectors by the sine of the angle between their positive directions, that is, equal to the magnitude of either vector by the normal component of the other. This product is designated by the symbol $\mathbf{P} \times \mathbf{Q}$, the cross between the two vectors representing the sine of the angle between them.

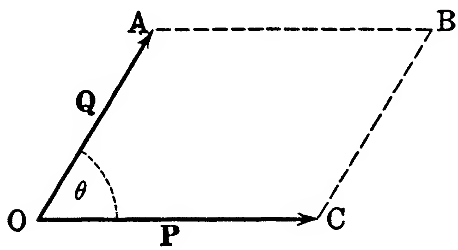


FIG. 119

If, then, the two vectors **P** and **Q** are oriented as in Fig. 119 the vector $\mathbf{P} \times \mathbf{Q}$ is directed out from the paper and has the magnitude $PQ \sin \theta$. On the other hand, $\mathbf{Q} \times \mathbf{P}$ is a vector of the same magnitude directed into the paper. Hence

$$\mathbf{Q} \times \mathbf{P} = -\mathbf{P} \times \mathbf{Q}. \tag{68-7}$$

Whenever, then, we interchange two vectors appearing in a vector product we must change the sign of the product.

Since the area of the parallelogram $OABC$ is $PQ \sin \theta$, the vector product $\mathbf{P} \times \mathbf{Q}$ has a magnitude equal to the area of the parallelogram of which the vectors **P** and **Q** are the sides. Moreover, as the vector product of **P** and **Q** is directed along the normal to the parallelogram, the vector representing the surface is $\mathbf{P} \times \mathbf{Q}$ or $\mathbf{Q} \times \mathbf{P}$ according as the one or the other face of the parallelogram is chosen as positive.

The torque **L** exerted by a force **F** about a point O provides a simple example of the vector product. For if **r** is the position vector of the point of application of the force relative to O,

$$\mathbf{L} = \mathbf{r} \times \mathbf{F}, \tag{68-8}$$

the vector product having the magnitude of the moment of the force and the direction of the axis about which it tends to produce rotation.

We shall find it very convenient to use the short-hand notation which we have developed here. In so far as the major portion of the text is concerned no more vector analysis will be used than that which has been presented up to this point. All that the reader needs to remember is that $\mathbf{P} \cdot \mathbf{Q}$ represents the scalar $PQ \cos \theta$, where θ is the angle between the vectors $\mathbf{P}$ and $\mathbf{Q}$, and that $\mathbf{P} \times \mathbf{Q}$ represents a vector perpendicular to the plane of $\mathbf{P}$ and $\mathbf{Q}$ in the sense of advance of a right-handed screw rotated from $\mathbf{P}$ to $\mathbf{Q}$ having a magnitude equal to $PQ \sin \theta$. To solve a few of the problems and to develop part of the theory given in Chapter XVI, however, it is necessary to express the scalar and vector products in terms of the components of the vectors involved. Therefore we shall now investigate the expansions of the two products.

Expansion of the Scalar Product. — First we must note that the distributive law of multiplication holds for the scalar product. This law states that

$$(\mathbf{P} + \mathbf{Q}) \cdot \mathbf{R} = \mathbf{P} \cdot \mathbf{R} + \mathbf{Q} \cdot \mathbf{R}, \tag{68–9}$$

which is evidently true since it amounts to nothing more than the statement that the component of $\mathbf{P} + \mathbf{Q}$ in the direction of $\mathbf{R}$ is equal to the sum of the component of $\mathbf{P}$ in the direction of $\mathbf{R}$ and the component of $\mathbf{Q}$ in the direction of $\mathbf{R}$.

Having shown that the distributive law is valid for the scalar product of the sum of two vectors by a third, it follows at once that it holds for the sum of any number of vectors. For example,

$$\begin{aligned}(\mathbf{P} + \mathbf{Q} + \mathbf{R})\cdot\mathbf{S} = \{(\mathbf{P} + \mathbf{Q}) + \mathbf{R}\}\cdot\mathbf{S} &= (\mathbf{P} + \mathbf{Q})\cdot\mathbf{S} + \mathbf{R}\cdot\mathbf{S}\\ &= \mathbf{P}\cdot\mathbf{S} + \mathbf{Q}\cdot\mathbf{S} + \mathbf{R}\cdot\mathbf{S},\end{aligned}$$

and similarly,

$$\begin{aligned}(\mathbf{P} + \mathbf{Q} + \mathbf{R})\cdot(\mathbf{S} + \mathbf{T} + \mathbf{U}) = {}&\mathbf{P}\cdot\mathbf{S} + \mathbf{P}\cdot\mathbf{T} + \mathbf{P}\cdot\mathbf{U}\\ &+ \mathbf{Q}\cdot\mathbf{S} + \mathbf{Q}\cdot\mathbf{T} + \mathbf{Q}\cdot\mathbf{U} + \mathbf{R}\cdot\mathbf{S} + \mathbf{R}\cdot\mathbf{T} + \mathbf{R}\cdot\mathbf{U}.\end{aligned}$$

The scalar products of the unit vectors $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ are

$$\boldsymbol{i}\cdot\boldsymbol{i} = \boldsymbol{j}\cdot\boldsymbol{j} = \boldsymbol{k}\cdot\boldsymbol{k} = 1,$$

since the angle involved in each is 0, and

$$\boldsymbol{i}\cdot\boldsymbol{j} = \boldsymbol{j}\cdot\boldsymbol{k} = \boldsymbol{k}\cdot\boldsymbol{i} = 0,$$

as the angle is $\pi/2$. Therefore the scalar product of the vectors **P** and **Q** given in (68–1) and (68–2) is

$$\begin{aligned}\mathbf{P}\cdot\mathbf{Q} &= (\boldsymbol{i}P_x + \boldsymbol{j}P_y + \boldsymbol{k}P_z)\cdot(\boldsymbol{i}Q_x + \boldsymbol{j}Q_y + \boldsymbol{k}Q_z)\\ &= \boldsymbol{i}\cdot\boldsymbol{i}P_xQ_x + \boldsymbol{i}\cdot\boldsymbol{j}P_xQ_y + \boldsymbol{i}\cdot\boldsymbol{k}P_xQ_z + \text{etc.}\\ &= P_xQ_x + P_yQ_y + P_zQ_z. \end{aligned} \qquad (68\text{–}10)$$

Thus the work done by a force

$$\mathbf{F} = \boldsymbol{i}F_x + \boldsymbol{j}F_y + \boldsymbol{k}F_z$$

when the body on which it is acting undergoes a displacement

$$d\mathbf{l} = \boldsymbol{i}dx + \boldsymbol{j}dy + \boldsymbol{k}dz$$

is

$$\mathbf{F}\cdot d\mathbf{l} = F_xdx + F_ydy + F_zdz,$$

each term on the right representing the work done by one component of the force.

Expansion of the Vector Product. — The distributive law holds also for the vector product, that is,

$$(\mathbf{P} + \mathbf{Q}) \times \mathbf{R} = \mathbf{P} \times \mathbf{R} + \mathbf{Q} \times \mathbf{R}. \qquad (68\text{–}11)$$

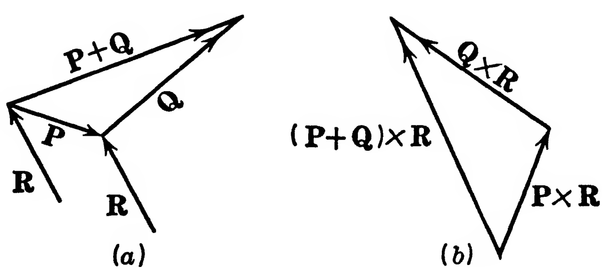


FIG. 120

To prove this, consider the vectors **P**, **Q**, **P** + **Q**, **R** as illustrated in Fig. 120a. Project the triangle formed by the vectors **P**, **Q**, **P** + **Q** on a plane perpendicular to **R**. The sides of the

projected triangle represent the components of the three vectors **P**, **Q**, **P** + **Q** normal to **R**. If, therefore, the projected triangle is rotated in its plane through a right angle and the length of each side increased in the ratio $R : 1$ the sides of the resulting triangle (Fig. 120b) will represent in magnitude and direction the vectors $\mathbf{P} \times \mathbf{R}$, $\mathbf{Q} \times \mathbf{R}$, $(\mathbf{P} + \mathbf{Q}) \times \mathbf{R}$. These three vectors, then, form a closed triangle, showing that the third is the sum of the first and second.

In developing the expansion of the vector product it is important that the axes employed should be right-handed, that is, that the Z axis should have the direction of advance of a right-handed screw rotated from X to Y through the right angle between them. If this condition is satisfied the vector products of the unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ are

$$\boldsymbol{i} \times \boldsymbol{i} = \boldsymbol{j} \times \boldsymbol{j} = \boldsymbol{k} \times \boldsymbol{k} = 0,$$
$$\boldsymbol{i} \times \boldsymbol{j} = \boldsymbol{k}, \qquad \boldsymbol{j} \times \boldsymbol{k} = \boldsymbol{i}, \qquad \boldsymbol{k} \times \boldsymbol{i} = \boldsymbol{j}.$$

Consequently

$$\begin{aligned}\mathbf{P} \times \mathbf{Q} &= (\boldsymbol{i}P_x + \boldsymbol{j}P_y + \boldsymbol{k}P_z) \times (\boldsymbol{i}Q_x + \boldsymbol{j}Q_y + \boldsymbol{k}Q_z) \\ &= \boldsymbol{i} \times \boldsymbol{i}P_xQ_x + \boldsymbol{i} \times \boldsymbol{j}P_xQ_y + \boldsymbol{i} \times \boldsymbol{k}P_xQ_z + \text{etc.} \\ &= \boldsymbol{i}(P_yQ_z - P_zQ_y) + \boldsymbol{j}(P_zQ_x - P_xQ_z) \\ &\qquad + \boldsymbol{k}(P_xQ_y - P_yQ_x). \quad (68\text{–}12)\end{aligned}$$

This expansion can be written very compactly in determinant notation as

$$\mathbf{P} \times \mathbf{Q} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ P_x & P_y & P_z \\ Q_x & Q_y & Q_z \end{vmatrix}. \qquad (68\text{–}13)$$

As an example, suppose we wish to write down the components of the torque given by (68–8). As

$$\mathbf{r} = \boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z,$$

the components of **L** are

$$\begin{aligned}L_x &= yF_z - zF_y, \\ L_y &= zF_x - xF_z, \\ L_z &= xF_y - yF_x.\end{aligned}$$

Problem 68a. A rigid body has an angular velocity $\boldsymbol{\omega}$ about an axis through the origin. Show that the linear velocity of a point whose position vector is $\mathbf{r}$ is $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$.

Problem 68b. Show that the scalar $(\mathbf{P} \times \mathbf{Q})\cdot\mathbf{R}$ represents the volume of the parallelopiped of which $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ are the edges. Show therefore that $(\mathbf{P} \times \mathbf{Q})\cdot\mathbf{R} = \mathbf{P}\cdot(\mathbf{Q} \times \mathbf{R})$, that is, the dot and the cross are interchangeable.

69. Ampère's Law. — As was pointed out in article 48 an electric current along a conducting wire is attributed to a drift of electrons in the direction opposite to that of the impressed electric force. If a point charge q is at rest in the observer's inertial system the only field to which it gives rise is the electric field directed along the radius vector,

$$\mathbf{E} = \frac{q}{r^3}\mathbf{r}. \tag{69–1}$$

In 1819, however, the Danish physicist Oersted discovered that a current exerts a torque on a magnet held near it which tends to turn the magnet until its axis is at right angles to the wire. In the following year Ampère showed that currents also exert forces on one another and succeeded in formulating analytically the force or torque exerted by one current circuit on another. Although Oersted's and Ampère's experiments were confined to conduction currents in wires, Rowland showed in 1876 that a moving electrified body produces the same effect on a magnet as a current in a wire.

We conclude, then, that a charge moving relative to the observer's inertial system produces a magnetic field as well as an electric field. The magnitude of the magnetic field is proportional to the speed of the charge relative to the observer, provided the speed is small compared to the velocity of light. An observer traveling with a moving charge, for instance, detects no magnetic field. In other words, a moving charge exerts no force or torque on a magnet which is moving along with it, but only on a magnet which does not partake of its motion.

Provided the velocity $\mathbf{v}$ of a moving point charge q is small compared to that of light, the magnetic intensity at a point P

(Fig. 121) at a distance $\mathbf{r}$ from the charge is found to be

$$\mathbf{H} = \frac{q\mathbf{v} \times \mathbf{r}}{r^3}. \tag{69–2}$$

This magnetic field is perpendicular to the plane of $\mathbf{v}$ and $\mathbf{r}$ in the direction of advance of a right-handed screw rotated from the first to the second of these vectors. At the point P, therefore, $\mathbf{H}$ is directed toward the reader. Its magnitude is equal to the product of q, v and $\sin\theta$ divided by the square of the distance r.

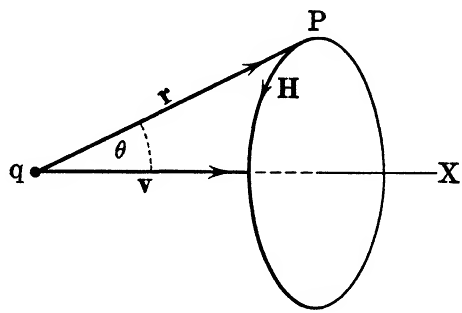


FIG. 121

If we take instead of P other points on the circumference of a circle passing through P and lying in a plane perpendicular to $\mathbf{v}$ with its center on the X axis, we note that the magnitude of $\mathbf{H}$ remains unaltered. Its direction, however, changes so as to be everywhere tangent to the circle. The lines of magnetic force, therefore, are circles in planes perpendicular to the velocity with their centers on a line through the charge in the direction of the velocity. Moreover, the sense in which these lines of force are described is that of the rotation of a right-handed screw advancing in the direction of the velocity. Finally, if we consider different points on the surface of a sphere with center at q, we note that H vanishes at the poles $\theta = 0$ and $\theta = \pi$ and is a maximum along the equator $\theta = \pi/2$.

In equation (69–1) all the quantities involved are measured in electrostatic units. If we should express q in (69–2) in electrostatic units, however, we should find that the magnetic intensity calculated from this equation did not represent the force on a unit pole as defined in the chapter on magnetostatics, but the force on a very much larger pole. Therefore, as it is convenient

to continue to measure magnetic fields in eléctromagnetic units or gauss, we must introduce a new unit of charge which we may define as that charge which, when moving with a velocity of 1 cm/sec, produces a magnetic field of 1 gauss at a point in the equatorial plane at a distance of 1 cm from the charge. This unit, known as the *electromagnetic unit of charge*, is the one which we shall employ throughout the subject of current electricity. Later we shall see that this unit is larger than the electrostatic unit of charge in the ratio $3(10)^{10} : 1$.

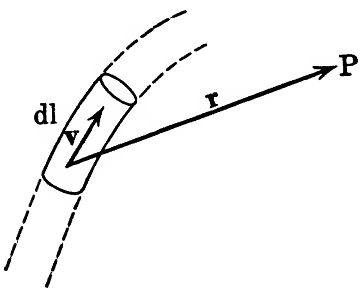


FIG. 122

Consider, now, a small section dl (Fig. 122) of a wire of cross-section A carrying a current $\mathbf{i}$. If n is the number of free electrons per unit volume, e the charge of each and $\mathbf{v}$ their drift velocity,

$$\mathbf{i} = Ane\mathbf{v}. \tag{69–3}$$

As we are measuring the charge e in electromagnetic units this equation gives the current $\mathbf{i}$ in electromagnetic units. It may be considered as defining the electromagnetic unit of current.

The total moving charge in the element under consideration is

$$q = Anedl. \tag{69–4}$$

Substituting this in (69–2), the field at P due to the length dl of the circuit is found to be

$$d\mathbf{H} = \frac{(Ane\mathbf{v}) \times \mathbf{r}}{r^3} dl$$

$$= \frac{\mathbf{i} \times \mathbf{r}}{r^3} dl, \tag{69–5}$$

by virtue of (69–3). As $d\mathbf{l}$ may be considered to be a vector in the direction of the current this may be written equally well

$$d\mathbf{H} = i\frac{d\mathbf{l} \times \mathbf{r}}{r^3}. \tag{69–6}$$

This is Ampère's law for the magnetic field due to a current element. In the case of a current in a wire, the electric field due to the moving electrons is annulled by the opposite electric field produced by the stationary atomic nuclei, so that the only electric field inside the wire is the impressed field along its length which is responsible for the current.

Equation (69–5) specifies the magnetic field produced by a current element. With the aid of the law of action and reaction we can deduce from this equation the force exerted on a current element by an external magnetic field. To do so, consider a pole m placed at P (Fig. 122). The force exerted on the pole by the current element is

$$d\mathbf{F}_m = m d\mathbf{H} = \mathbf{i} \times \left(\frac{m\mathbf{r}}{r^3} \right) dl, \qquad (69\text{–}7)$$

by (69–5). According to the law of action and reaction the force exerted by m on the current element is equal and opposite to this. Now the magnetic field at the current element due to the pole is

$$\mathbf{H} = -\frac{m\mathbf{r}}{r^3}.$$

Consequently the force $d\mathbf{F}$ on the current element is

$$d\mathbf{F} = \mathbf{i} \times \mathbf{H} dl. \qquad (69\text{–}8)$$

As this expression contains no reference to the pole m, it represents the force due to an external field $\mathbf{H}$, no matter what the origin of the field may be. Dividing by dl, the force per unit length of the circuit is

$$\mathbf{F}_l = \mathbf{i} \times \mathbf{H}. \qquad (69\text{–}9)$$

Again, if we replace $\mathbf{i}$ by $\mathbf{j}A$ in (69–8),

$$d\mathbf{F} = \mathbf{j} \times \mathbf{H} A dl,$$

and the force per unit volume of the current is

$$\mathbf{F}_\tau = \mathbf{j} \times \mathbf{H}. \qquad (69\text{–}10)$$

To find the force on a moving charge due to a magnetic field

through which it may be passing, replace **i** in (69–8) by the right-hand member of (69–3) and then use (69–4). This gives

$$\mathbf{F} = q\mathbf{v} \times \mathbf{H}. \tag{69–11}$$

We must consider next what modifications, if any, are required in the formulas of this article when the currents under consideration are surrounded by a magnetic medium. First consider Ampère's law (69–5) or (69–6) for the field produced by a current element. As the lines of force in the neighborhood of the current are closed curves surrounding the circuit, the intensity of magnetization in the medium has no component normal to the surface of the cavity in which the circuit lies. Therefore the field due to the current is not altered by the induction of magnetic charges on the surface of the medium adjacent to the circuit, and equations (69–5) and (69–6) are valid whether the current is surrounded by a magnetic medium or not. The same statement applies to the expression (69–2) for the magnetic intensity produced by a moving charge.

When we come to consider the force (69–8) on a current element due to an external magnetic field we must distinguish several cases. First consider the case, often encountered in practice, of a non-magnetic wire CD (Fig. 123) carrying a current **i** which is free to move in a slot in a solid magnetized medium, such as iron. If we understand by **H** the magnetic intensity in the slot, (69–8) remains unaltered. But it is often more convenient to express the force in terms of the field in the iron. At first let us limit ourselves to the case where **B** in the iron lies in the plane of the wire and the normal n to the face of the slot.

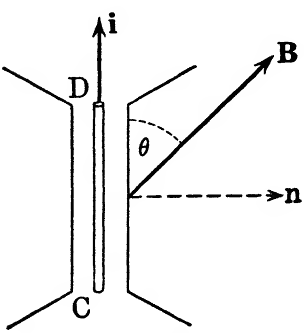


FIG. 123

Since the force on the current is determined solely by the component H_n of the field in the slot at right angles to the circuit, the force on a length dl of the current is

$$dF = iH_n dl$$

into the paper. But, as the normal component of the magnetic induction is continuous across the surface of the slot, $H_n = B \sin \theta$, and we may write

$$dF = iB \sin \theta dl,$$

or, in vector notation,

$$d\mathbf{F} = \mathbf{i} \times \mathbf{B}dl. \tag{69–12}$$

If, now, the field has a component at right angles to the plane of the figure, this component of the magnetic intensity, being tangent to the face of the slot, has the same value in the iron as in the slot. So far as the force due to it is concerned, we must replace **B** by **H** in (69–12). But the force on the current due to the component of the field under consideration is along the normal n. If, then, we are concerned only with the component of the force parallel to the face of the slot, formula (69–12) will always yield the correct result, **B** representing the magnetic induction in the iron.

Next we have the case of a current in a non-magnetic wire which is immersed in a paramagnetic fluid. The detailed treatment of this case is beyond the scope of this book. It may be shown, however, that formula (69–12) holds, **B** representing the mean magnetic induction in the medium before the wire is placed in it.

Finally we have the case of a current passing through a medium which is at the same time conducting and magnetic, such as a current flowing through a magnet. This case will be treated in full in article 75, where it will be shown that (69–12) is the correct expression for the force.

Problem 69a. A current element idl is located at the origin, the current having the direction of the Z axis. What are the components of H at a point $P(x, y, z)$? Ans. $-\frac{iydl}{r^3}, \frac{ixdl}{r^3}, 0.$

Problem 69b. A current of density j lies in a magnetic field H parallel to the X axis. If the direction cosines of the current are l, m, n, find the components of the force on it per unit volume. Ans. $0, njH, -mjH.$

70. Magnetic Fields of Simple Circuits. — The magnetic field at the center O (Fig. 124) of a circular circuit of radius a consisting of a single turn of wire is easily calculated. The magnetic intensity due to the current element AB of length dl is evidently perpendicular to the plane of the circle toward the reader. From (69–5) its magnitude is

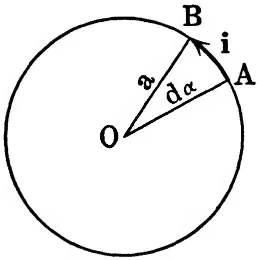


FIG. 124

$$dH = \frac{i}{a^2} dl,$$

since the angle between AB and AO is $\pi/2$. Now $dl = a d\alpha$, and consequently the total field at O is

$$H = \frac{i}{a}\int_0^{2\pi} d\alpha = \frac{2\pi i}{a}, \qquad (70\text{–}1)$$

directed out from the paper. The lines of force in the field of a circular current are illustrated in Fig. 125. If the circuit con-

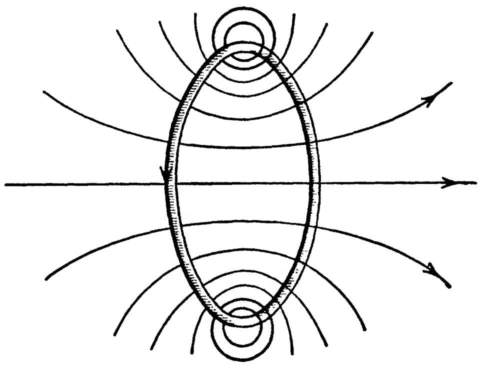


FIG. 125

sists of n turns of wire lying close together, each turn gives rise to a field of this magnitude and we have altogether

$$H = \frac{2\pi n i}{a}. \qquad (70\text{–}2)$$

Next we shall calculate the field due to a circular circuit of one turn at a point P (Fig. 126) on the axis of the circle at a distance x from the center O. First consider the field $d\mathbf{H}$ due to the current element AB. It is directed at right angles to i and the radius vector AP, and as the angle between these is $\pi/2$ its magnitude is

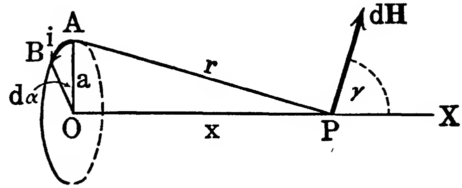


FIG. 126

$$dH = \frac{iad\alpha}{r^2}.$$

Two diametrically opposite current elements give rise to fields whose components normal to OX annul each other. Therefore the resultant field is along the X axis and equal to

$$H = \frac{ia}{r^2}\left[\int_0^{2\pi} d\alpha\right]\cos\gamma = \frac{2\pi i a^2}{r^3}$$

$$= \frac{2\pi i a^2}{(a^2 + x^2)^{3/2}}. \qquad (70\text{–}3)$$

Note that H is in the direction of advance of a right-handed screw rotated in the same sense as the current.

We are ready now to compute the field on the axis of a straight solenoid or coil of circular cross-section such as is shown

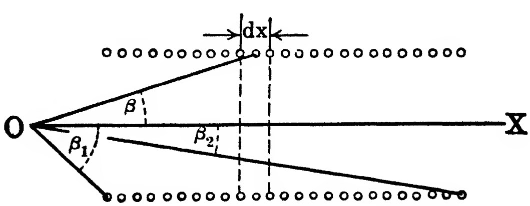


FIG. 127

in longitudinal section in Fig. 127. If n_1 is the number of turns per unit length, the number of turns in a length dx is n_1dx, and

the field at O due to an element dx of a solenoid of radius a is

$$dH = \frac{2\pi i a^2 n_1 dx}{(a^2 + x^2)^{3/2}}$$

from (70–3). The magnetic intensity is directed along the axis OX of the solenoid in the sense of advance of a right-handed screw rotated in the same sense as the current.

Now

$$x = a \cot \beta, \qquad dx = -a \csc^2 \beta d\beta,$$

and hence

$$\begin{aligned} H &= -2\pi n_1 i \int_{\beta_1}^{\beta_2} \sin \beta d\beta \\ &= 2\pi n_1 i (\cos \beta_2 - \cos \beta_1). \end{aligned} \tag{70–4}$$

Near the center of a solenoid of length very great compared to its radius, $\cos \beta_1 = -1$, $\cos \beta_2 = 1$ and

$$H = 4\pi n_1 i. \tag{70–5}$$

At the end of the solenoid $\cos \beta_1 = 0$, $\cos \beta_2 = 1$ and

$$H = 2\pi n_1 i. \tag{70–6}$$

Therefore half the lines of force passing through the central section of a long solenoid pass out through the sides before reaching the end. Comparing with the results of problem 43*b* we see that the H field inside a long solenoid is identical with the B field inside a long uniformly magnetized bar magnet. Figure 128 shows the distribution of the lines of force.

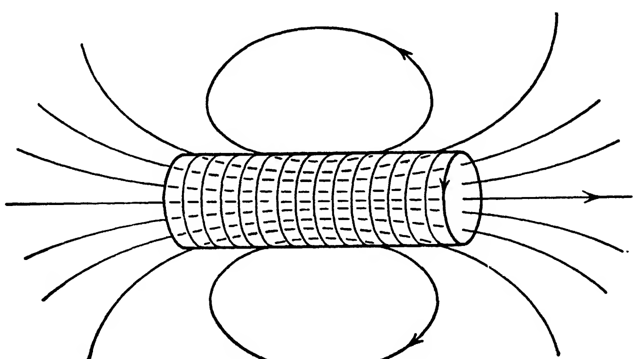


FIG. 128

Next we shall calculate the field at a point O (Fig. 129) due to a straight current AB parallel to the Y axis. As $\sin\theta = \cos\beta$ the magnetic intensity at O due to the current element idy is

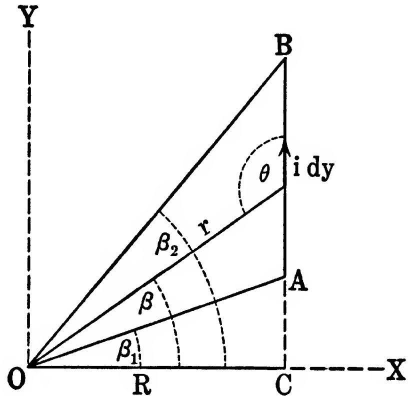


FIG. 129

$$dH = \frac{idy\cos\beta}{r^2},$$

directed out from the paper. If R is the perpendicular distance OC of the current from O,

$$\begin{aligned} y &= R\tan\beta, \\ dy &= R\sec^2\beta d\beta \\ &= r\sec\beta d\beta. \end{aligned}$$

Therefore $dy\cos\beta = rd\beta$ and

$$dH = \frac{id\beta}{r} = \frac{i\cos\beta d\beta}{R}.$$

Integrating,

$$H = \frac{i}{R}(\sin\beta_2 - \sin\beta_1). \qquad (70\text{–}7)$$

If the straight current is infinitely long, $\sin\beta_1 = -1$, $\sin\beta_2 = 1$ and

$$H = \frac{2i}{R}, \qquad (70\text{–}8)$$

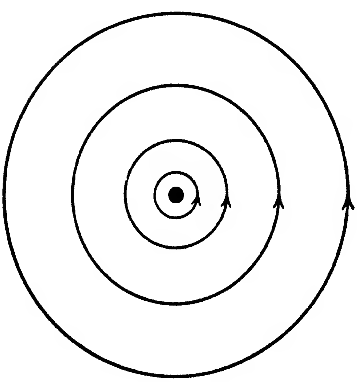


FIG. 130

an expression known as *the law of Biot and Savart.* Evidently the lines of force are circles in planes perpendicular to the current with centers on the latter. They are depicted in Fig. 130, the current being directed toward the reader. The sense in which

they are described is that of rotation of a right-handed screw which advances in the direction of the current. We can use (70–7) to compute the field at any point due to a circuit in the shape of a polygon, whether plane or not.

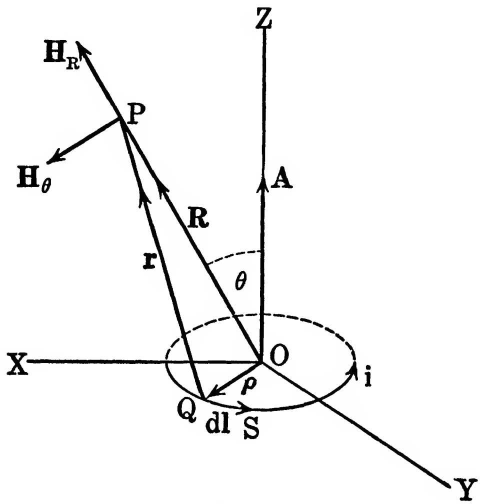


FIG. 131

Finally we shall evaluate the field due to a small plane circuit of any shape (Fig. 131) at a point P whose distance from the circuit is large compared to the linear dimensions of the latter. Take origin at some point O inside the circuit and orient axes so that the Z axis is along the positive normal to the plane of the circuit and OP lies in the XZ plane. If $\boldsymbol{\rho}$ is the position vector of the current element QS relative to O the vector distance QP is

$$\mathbf{r} = \mathbf{R} - \boldsymbol{\rho}.$$

The square of the distance QP is

$$r^2 = (\mathbf{R} - \boldsymbol{\rho})\cdot(\mathbf{R} - \boldsymbol{\rho}) = R^2 - 2\mathbf{R}\cdot\boldsymbol{\rho} + \rho^2,$$

where $\rho^2 = x^2 + y^2$. Now $\mathbf{R}\cdot\boldsymbol{\rho}$ is the projection of $\boldsymbol{\rho}$ on $\mathbf{R}$ multiplied by R, that is, $Rx \sin\theta$. Therefore,

$$r^2 = R^2 - 2Rx \sin\theta + \rho^2.$$

As x and y are small compared to R we can neglect terms involving the squares of these coordinates throughout our analysis. Hence

$$\frac{1}{r^3} = \frac{1}{R^3\left(1 - 2\frac{x}{R}\sin\theta\right)^{3/2}} = \frac{1}{R^3}\left(1 + 3\frac{x}{R}\sin\theta\right)$$

by the binomial theorem. Therefore Ampère's law (69–6) gives us for the field due to a current element of length dl

$$d\mathbf{H} = \frac{i}{R^3}\left(1 + 3\frac{x}{R}\sin\theta\right)(d\mathbf{l} \times \mathbf{R} - d\mathbf{l} \times \boldsymbol{\rho})$$

$$= \frac{i}{R^3}\left(d\mathbf{l} \times \mathbf{R} - d\mathbf{l} \times \boldsymbol{\rho} + 3\frac{x}{R}\sin\theta d\mathbf{l} \times \mathbf{R}\right).$$

Summing up over all the current elements in the circuit,

$$\mathbf{H} = \frac{i}{R^3}\left(\oint d\mathbf{l}\right) \times \mathbf{R} - \frac{i}{R^3}\oint (d\mathbf{l} \times \boldsymbol{\rho}) + 3\frac{i}{R^4}\sin\theta\left(\oint x d\mathbf{l}\right) \times \mathbf{R}.$$

Now $\oint d\mathbf{l} = 0$ as the circuit is a closed curve. The vector representative of the triangle OQS is $\frac{1}{2}\boldsymbol{\rho} \times d\mathbf{l}$ since the area of this triangle is half that of the parallelogram of which $\boldsymbol{\rho}$ and $d\mathbf{l}$ are the sides. Therefore $\oint \boldsymbol{\rho} \times d\mathbf{l}$ is twice the vector area $\mathbf{A}$ of the circuit, this vector being directed perpendicular to the circuit along the Z axis. Finally, as $d\mathbf{l} = \boldsymbol{i}dx + \boldsymbol{j}dy$, where $\boldsymbol{i}$ and $\boldsymbol{j}$ are unit vectors along the X and Y axes respectively,

$$\oint x d\mathbf{l} = \boldsymbol{i}\oint x dx + \boldsymbol{j}\oint x dy = \boldsymbol{i}\left|\frac{x^2}{2}\right|_x^x + \boldsymbol{j}A = \boldsymbol{j}A,$$

for the first integral vanishes since the upper and lower limits are the same and the second integral is the familiar expression for the area under a curve. Therefore

$$\mathbf{H} = 2\frac{i\mathbf{A}}{R^3} + 3\frac{iA\sin\theta}{R^4}\boldsymbol{j} \times \mathbf{R}.$$

Taking components along R and at right angles thereto in

the direction of increasing θ,

$$\left.\begin{aligned} H_R &= \frac{2(iA)}{R^3}\cos\theta, \\ H_\theta &= \frac{(iA)}{R^3}\sin\theta. \end{aligned}\right\} \quad (70\text{–}9)$$

Comparing these expressions with (40–2) we see that the field at a distance from a current circuit large compared to its linear dimensions is identical with that produced by a small magnet of moment equal to the product of the current by the area of the circuit, the magnetic axis of the equivalent magnet being normal to the plane of the circuit in the sense of advance of a right-handed screw rotated in the sense in which the current is flowing. For this reason the product of the current by the area of a plane circuit is known as the *magnetic moment* of the circuit.

As a small current circuit produces the same field as a small magnet we can use the law of action and reaction to prove that it must experience the same force or torque as a small magnet of moment iA when placed in an external magnetostatic field.

Problem 70a. A circuit is in the form of a square of side l. Find the magnetic field at its center. Ans. $\dfrac{8\sqrt{2}i}{l}$.

Problem 70b. A circuit has the form of a regular polygon of n sides inscribed in a circle of radius a. Find H at the center and show that the result goes over into (70–1) as n is indefinitely increased. Ans. $\dfrac{2ni}{a}\tan\dfrac{\pi}{n}$.

Problem 70c. Find H due to the circuit of problem 70*a* at a point on the axis of the square at a distance d from its center.

$$\text{Ans. } \frac{2il^2}{\left(d^2 + \dfrac{l^2}{4}\right)\sqrt{d^2 + \dfrac{l^2}{2}}}.$$

Problem 70d. Two long parallel wires a distance $2d$ apart carry equal but oppositely directed currents. Find H at a point in the plane of the wires distant R from a line half-way between the two.

$$\text{Ans. } \frac{4id}{R^2 - d^2}.$$

Problem 70e. The length of a solenoid is ten times its radius. What is the percentage error in using (70–5) to calculate the field at the center? Ans. 2%.

Problem 70f. The windings of an actual solenoid are in the form of a helix rather than a succession of circles such as was assumed in the previous article. Show, however, that (70–4) for the axial field is valid also for helical windings.

Problem 70g. A solenoid is 2 cm in diameter and 20 cm long. Plot the axial field against the distance from the center of the solenoid.

Problem 70h. A solenoid of n_1 turns per unit length is wound in a number of layers extending from a distance a from the axis to a distance b Find the axial field at a point in the interior distant l_1 and l_2 from the two ends.

$$\text{Ans. } \frac{2\pi n_1 i}{b-a}\left\{ l_1 \log_e \frac{\sqrt{l_1^2+b^2}+b}{\sqrt{l_1^2+a^2}+a} + l_2 \log_e \frac{\sqrt{l_2^2+b^2}+b}{\sqrt{l_2^2+a^2}+a} \right\}.$$

Problem 70i. A circular coil of n turns is wound in a rectangular channel of width $2d$ and depth $2h$. If a is the mean radius of the coil, find the field at its center.

$$\text{Ans. } \frac{2\pi n i}{a}\left\{ 1 + \frac{2h^2 - 3d^2}{6a^2} + \cdots \right\}.$$

Problem 70j. A unit positive pole is carried along the axis of a circular current from infinity on the positive side of the circuit to infinity on the negative side. Show that the work done is $4\pi i$.

71. Magnetic Shell Equivalent to Current Circuit. — We shall show now that the magnetic field produced by a current circuit is identical with that produced by a magnetic shell of constant strength equal to the current whose periphery coincides with the circuit. This statement is limited, however, to the region outside the equivalent shell, for at points inside the shell the field of the latter differs both in magnitude and direction from that of the circuit.

We shall start by considering a magnetic shell of constant strength Φ in the shape of the plane figure $ABCD$ (Fig. 132) of infinitesimal width extending from AB an infinite distance to the left. Let the sides CB and DA make the angle $d\alpha$ at P and take the front face of the shell as positive. We shall calculate the magnetic intensity at P due to the shell. The magnetic moment of the section of the shell included between the arcs of

radii p and $p + dp$ described about P is $\Phi p d\alpha dp$ directed outward from the paper, and from (40–2) the field at P due to this

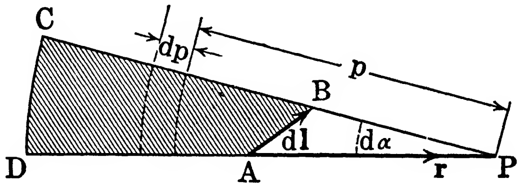


FIG. 132

element of the shell is

$$d^2H = \frac{\Phi p d\alpha dp}{p^3} = \Phi d\alpha \frac{dp}{p^2}$$

directed inwards, as $\theta = \pi/2$ in the formula. Integrating with respect to p from r to ∞ the field at P due to the entire shell is found to be

$$dH = \Phi d\alpha \int_r^\infty \frac{dp}{p^2} = \frac{\Phi d\alpha}{r} = \frac{\Phi r^2 d\alpha}{r^3}.$$

Now $r^2 d\alpha$ is twice the area of the triangle ABP. But if $\mathbf{r}$ is the vector distance AP and $d\mathbf{l}$ the vector AB then $d\mathbf{l} \times \mathbf{r}$ is a vector directed into the paper having a magnitude equal to the area of the parallelogram of which AP and AB are the sides, that is, equal to twice the area of the triangle ABP. Therefore the magnetic intensity at the point P is given in magnitude and direction by the vector expression

$$d\mathbf{H} = \Phi \frac{d\mathbf{l} \times \mathbf{r}}{r^3}. \qquad (71\text{–}1)$$

Comparing with Ampère's law (69–6) we see that this expression is identical with that for the field due to a current element $id\mathbf{l}$ along AB provided the current is equal to the strength Φ of the shell. Hence we conclude that the field at any point P due to a current element is equal to that of a long narrow magnetic shell of constant strength equal to the current extending, between two straight lines drawn from P through the ex-

tremities of the current element, from the current element to infinity.

Now consider the circuit $ABCD$ (Fig. 133). To find the magnetic intensity at P due to the current we may replace the circuit by a conical magnetic shell of strength equal to the current having the point P as vertex and extending from the curve $ABCD$ to infinity. As here described the shell is open at infinity. We

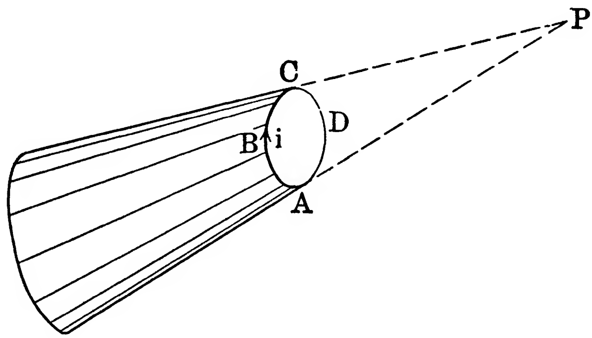


FIG. 133

may, however, close the farther end by means of a shell of the same strength in the form of a spherical cap without affecting the field at P. For the magnetic intensity at P due to an element of such a cap which subtends a solid angle $d\Omega$ at P is

$$dH = \frac{2\Phi r^2 d\Omega}{r^3} = \frac{2\Phi d\Omega}{r}$$

in the radial direction by (40–2). As r approaches infinity dH approaches zero, and the field of an infinitely distant spherical cap vanishes.

We have found now that the field at P due to the current is identical with that of a magnetic shell of constant strength equal to the current which is closed everywhere except over the curve $ABCD$. But we proved in article 41 that all shells of the same strength and the same periphery produce the same field at all outside points. Consequently we may replace the shell under consideration by any other shell of the same strength

whose periphery coincides with the circuit. Such a shell produces at all outside points a field identical with that of the equivalent current circuit. The positive face of the shell is related to the sense in which the current is flowing by the rule of article 8, that is, a right-handed screw, rotated in the sense of the current, advances from the negative toward the positive face of the equivalent shell.

The same conclusion may be reached by the following somewhat different line of reasoning. Let $ABCD$ (Fig. 134) be the circuit under consideration. Divide this circuit into a number of elementary circuits such as $abcd$ by means of the network shown in the figure and suppose that the same current i is flowing in each of these in the sense $ABCD$. Any one of the interior branches of the network, such as ab, is a part of two adjacent circuits and carries two equal and opposite currents the magnetic fields of which annul each other. Only those currents flowing in the peripheral branches are uncompensated. Therefore the resultant field is identical with that produced by a current i in $ABCD$.

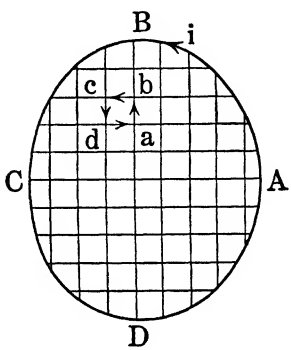


FIG. 134

Now it was shown in article 70 that an elementary current circuit produces the same magnetic field as a magnetic dipole parallel to the axis of the circuit of moment equal to the product of the current by the area of the circuit. The equivalent magnetic moment per unit area of the circuit is therefore equal to the current. Consequently the field produced by the network at exterior points is identical with that of a shell of strength i having the same periphery as the network, its positive face being uppermost.

As an illustration of the method which has just been developed for finding the field produced by a current circuit by substituting an equivalent magnetic shell, let us deduce H near the center of

a long solenoid of n_1 turns per unit length. Each turn of wire may be replaced by a plane shell of strength i. By making the thickness of each shell equal to $1/n_1$ the solenoid is replaced by a uniformly magnetized bar of the same outside dimensions having an intensity of magnetization $I = in_1$. At all outside points the field due to this equivalent magnet is identical with that of the solenoid. The equivalence, however, applies only to points outside the magnet. Therefore, in order to calculate the field inside the solenoid, we must leave a narrow transverse slit between the magnetic shells corresponding to the turns of wire on either side of the point P at which we wish to obtain the field strength. The poles on the ends of the bar being too far away to produce an appreciable effect at P, the field is due entirely to the magnetic charges on the surfaces of the slit. Consequently, if we make use of the result of problem 42*b*,

$$H = 4\pi I = 4\pi n_1 i,$$

in accord with (70–5). This method of treatment reveals one fact, however, which was not evident from the solution of article 70, namely, that the field is uniform through the entire cross-section of the solenoid provided the region under consideration is far away from the ends. Now the force on a unit positive pole placed inside a transverse slit in a magnetic medium is not H, but B, as was shown in problem 43*a*. Therefore the H field of a solenoid is identical with the B field of the equivalent magnet.

As shown in article 69, the magnetic intensity due to a current is unaltered when it is immersed in a permeable medium. The same is true of an infinitely thin magnetic shell, for, as the thickness of the shell approaches zero, the surface density of magnetic charge on the shell increases without limit, while the induced magnetization on the adjacent surface of the medium remains finite. Hence the contribution of the medium to the field vanishes with the thickness of the shell. We conclude, therefore, that the equivalent shell has a strength equal to the current even in a magnetic medium.

Problem 71a. Equal currents are flowing in opposite senses in two circuits not lying in the same plane. Show that the resultant magnetic field due to the two currents is the same as that due to a

single tubular magnetic shell extending from the one circuit to the other.

Problem 71b. A sphere of radius a carrying a uniform surface charge σ per unit area rotates about a diameter with constant angular velocity ω. Show that the magnetic field outside the sphere is the same as that of a stationary uniformly magnetized sphere of the same radius with intensity of magnetization $I = \sigma a \omega$.

Problem 71c. Show that the magnetic field inside the charged rotating sphere of 71*b* is a uniform field parallel to the axis of strength $(8\pi/3)\sigma a \omega$. (See art. 44.)

72. Circuital Form of Ampère's Law. — The results of the last article enable us to express Ampère's law in another form, which is often more useful than (69–5) or (69–6). Suppose the current circuit whose field we wish to investigate to be replaced by a very thin magnetic shell whose periphery coincides with the circuit and whose strength Φ is equal to the current i. It was shown in article 41 that the work done against the field of a shell of constant strength when a unit positive pole is taken from a point on the negative side around the edge of the shell to an opposite point on the positive side is $4\pi\Phi$. An equal amount of work is done by the field if the pole is allowed to move from the positive side of the shell around the edge to the negative side. But such a path encircles the current in the equivalent circuit in the sense of rotation of a right-handed screw which is advancing in the direction in which the current is flowing. Therefore the work done by the magnetic field produced by a current when a unit positive pole follows a path of any shape encircling the current in the sense of rotation of a right-handed screw advancing in the direction of the current is

$$W = 4\pi i. \qquad (72\text{–}1)$$

This is the circuital form of Ampère's law. In deducing it we have neglected the portion of the path of the pole lying inside the equivalent magnetic shell. If the field were due to a shell the work done in passing through the shell would be equal and opposite to that along the part of the path lying outside the shell, so that the net work done in carrying the pole around a

closed path would be zero. The field in the interior of the equivalent shell, however, is not the same as that due to the circuit. In the case of the latter the field is continuous all the way around. So, as the equivalent shell is extremely thin, the work done in carrying a pole from a point on one side to an opposite point on the other does not differ appreciably from the work done in completely encircling the current circuit.

If α is the angle between the magnetic intensity $\mathbf{H}$ produced by a current circuit and an element $d\mathbf{l}$ of a closed curve surrounding the current, the work done by the field when a unit positive pole is carried around the current is

$$W = \oint H \cos \alpha dl = \oint \mathbf{H} \cdot d\mathbf{l}. \tag{72–2}$$

This integral is known as a *magnetomotive force*. It is quite analogous to the *electromotive force* (48–5), namely,

$$\mathcal{E} = \oint E \cos \alpha dl = \oint \mathbf{E} \cdot d\mathbf{l}.$$

Comparison of (72–1) and (72–2) enables us to put Ampère's law in the form

$$\oint \mathbf{H} \cdot d\mathbf{l} = 4\pi i. \tag{72–3}$$

The closed path around which the integral is taken may be *any* closed curve arbitrarily chosen, i being the current passing through the curve in the direction of advance of a right-handed screw rotated in the sense in which the closed path is described. If the closed curve surrounds no current, the right-hand side of (72–3) is zero; if it encircles n wires each of which carries a current i the right-hand member becomes $4\pi ni$. Since the field produced by a current circuit is unaltered when the circuit is immersed in a magnetic medium, (72–3) holds no matter whether the circuit is located in empty space or is surrounded by a magnetic fluid or solid.

The circuital form of Ampère's law is especially useful in

calculating H in a field where symmetry shows that the magnetic intensity has the same magnitude at all points on a closed line of force which we choose for the path involved in (72–3). Consider, for instance, a straight wire AB (Fig. 135) of radius a along which a current i is flowing. We will provide a return circuit in the form of a cylindrical shell of radius b coaxial with the wire, the two being connected at their extremities by conducting disks in planes perpendicular to the wire. It is clear from symmetry that the lines of force are circles in planes perpendicular to the wire with their centers on its axis. Applying Ampère's law (72–3) to the circular line of force PQ of radius R lying between the wire and the cylindrical shell,

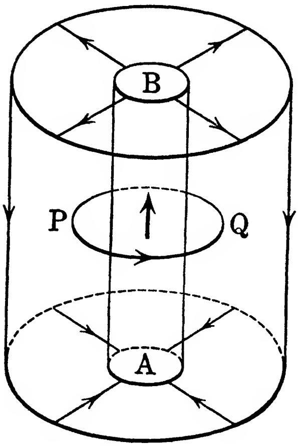


FIG. 135

$$2\pi RH = 4\pi i, \tag{72–4}$$

or

$$H = \frac{2i}{R}, \tag{72–5}$$

in agreement with the formula (70–8) for an infinitely long straight current.

If, however, R is less than a the path PQ lies inside the wire and encircles only a part of the current. If the current is steady and therefore distributed uniformly through the cross-section of the wire, the current surrounded is iR^2/a^2 and (72–4) becomes

$$2\pi RH = \frac{4\pi i R^2}{a^2},$$

leading to

$$H = \frac{2i}{a^2}R \tag{72–6}$$

at a point inside the wire.

On the other hand, if R is greater than b the path PQ encircles

the return circuit as well as the current in the wire. Therefore the net current passing through the closed curve along which

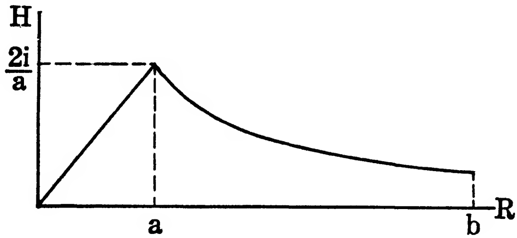


FIG. 136

we are integrating is zero and H vanishes. Consequently the field is confined entirely to the region inside the outer cylinder.

Plotting H against the distance R from the axis of the wire we get the graph shown in Fig. 136. By making b large enough we approach as closely as we choose to the ideal case of an infinitely long straight current.

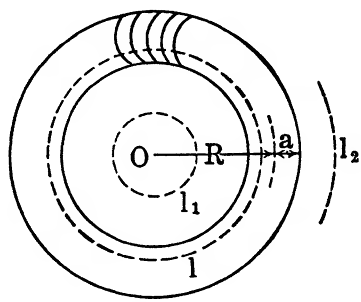


FIG. 137

Next we shall investigate the ring or toroidal solenoid (Fig. 137) wound uniformly with n turns. Evidently the lines of force are circles in planes parallel to the paper with centers on the axis through O. If H is the field along the path l lying inside the solenoid which encircles all n turns of wire, we have

$$Hl = 4\pi ni$$

from (72–3). Therefore

$$H = 4\pi \frac{n}{l} i \tag{72–7}$$

inside the coil. If the radius a of the solenoid is small compared to the radius R of its circular axis all paths such as l lying in its interior have nearly the same length and the field is nearly uniform. In this case (72–7) goes over into the formula (70–5)

for the field at a point on the axis of a straight solenoid far from its ends.

A path such as l_1 lying outside the solenoid surrounds no current and therefore H vanishes. On the other hand, a path such as l_2 encircles as many currents directed out from the paper as into the paper and hence H is zero here too. Consequently the field is confined entirely to the inside of the coil. These conclusions, including formula (72–7), apply no matter whether the cross-section of the ring is circular or not.

The straight solenoid previously discussed may be considered to be the limiting case of a ring solenoid when R becomes infinitely great, at least in so far as the field far from the ends is concerned. We can show very simply that the field inside such a solenoid near the middle must be uniform throughout its cross-section by applying (72–3) to the path $ABCD$ (Fig. 138). As the lines of force are parallel to the axis inside the solenoid and the field outside is negligibly small, the magnetomotive force reduces to the integral along AB. If the length of this line is l and n_1 is the number of turns per unit length,

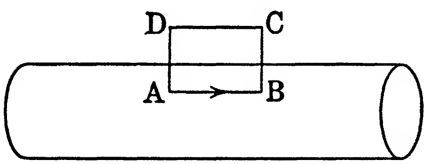


FIG. 138

$$Hl = 4\pi n_1 i l,$$

or

$$H = 4\pi n_1 i,$$

anywhere inside the solenoid far from the ends.

Problem 72a. A current flowing along a surface is known as a *current sheet.* Two parallel plane current sheets of infinite extent carry currents j per unit width in opposite senses. Find H (1) between the sheets, (2) outside the sheets. Ans. (1) $4\pi j$ at right angles to the current, (2) 0.

Problem 72b. Show that, if the lines of force of a magnetic field are straight and parallel in a region containing no current or magnetic charge, the magnetic intensity must remain constant in magnitude as we move in a direction at right angles to the field.

Problem 72c. Show that the field inside a ring solenoid is the same as that due to a straight current ni flowing along the axis of symmetry of the ring, n being the number of turns on the ring.

Problem 72d. Show that a many-valued potential of the form

$$V = -i(\Omega + 4\pi n)$$

exists in the case of the field of a current circuit, where Ω is the solid angle subtended by the circuit and n is an integer which increases by unity every time the circuit is linked.

Problem 72e. Although the components of the magnetic intensity produced by a current circuit cannot be expressed as derivatives of a single-valued scalar potential, they may be obtained from the *vector potential*

$$\mathbf{a} = \oint \frac{i d\mathbf{l}}{r}$$

integrated around the circuit. Referring to (69–6) show by differentiation under the sign of integration that

$$H_x = \frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z},$$

$$H_y = \frac{\partial a_x}{\partial z} - \frac{\partial a_z}{\partial x},$$

$$H_z = \frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y}.$$

73. Field of a Circular Current. — By the use of Ampère's law for the magnetic intensity due to a current element we have been able to calculate the field strength (70–3) on the axis of a circular current. Excepting such special cases, however, the field produced by a current circuit can be more readily computed by replacing the circuit by a magnetic shell which has the same periphery and a strength equal to the current, and then calculating the magnetic intensity due to the equivalent shell. As was noted in article 45 the magnetic potential in a magnetostatic field must satisfy Laplace's equation. Therefore the magnetic potential must be expressible as a solution or as the sum of a number of solutions of Laplace's equation so chosen as to satisfy the assigned boundary conditions. Once the potential has been found, the components of the magnetic intensity are obtained at once from (38–7) or (38–8).

We shall apply this method of solution to the calculation of the magnetic field produced by a circular current AB (Fig. 139) of radius a lying in a plane through the origin O at right angles to the X axis. Suppose that the current i enters the plane of the paper at A and emerges at B. Then the magnetic potential V

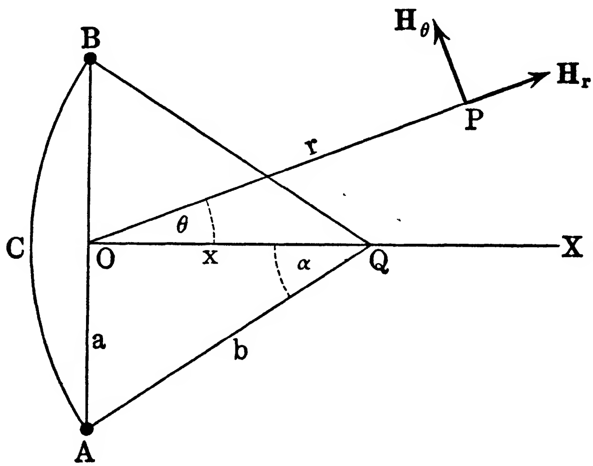


FIG. 139

at any point Q on the X axis may be calculated by replacing the circuit by a magnetic shell ACB of strength i in the shape of a spherical cap with Q as center, the concave face of the shell being positive.

If the angle which any radius vector drawn from Q to a point on the cap makes with the X axis is denoted by ϕ, the area of the cap is

$$S = \int_0^\alpha 2\pi b^2 \sin \phi d\phi = 2\pi b^2(1 - \cos \alpha),$$

and the solid angle subtended at Q by the cap is

$$\Omega = -\frac{S}{b^2} = -2\pi(1 - \cos \alpha) = -2\pi \left\{ 1 - \frac{x}{\sqrt{a^2 + x^2}} \right\}.$$

Therefore the potential at Q is

$$V = -i\Omega = 2\pi i \left\{ 1 - \frac{x}{\sqrt{a^2 + x^2}} \right\} \qquad (73\text{-}1)$$

from (41–1), since the point Q here is on the positive side of the shell. The field strength (70–3) at any point on the X axis is obtained at once from (73–1) by taking the negative of the derivative with respect to x.

We shall confine our attention to points at distances from O less than the radius a of the current. Then we can expand the radical in (73–1) in powers of x/a by the binomial theorem, getting

$$V = 2\pi i \left\{ 1 - \frac{x}{a} + \frac{1}{2}\frac{x^3}{a^3} + \cdots \right\}. \qquad (73\text{–}2)$$

This equation expresses the boundary condition which must be satisfied by the potential function. It is evident from symmetry that the potential at a point P off the X axis can be a function of r and θ only. Therefore it must consist of a series of zonal harmonics so chosen as to reduce to (73–2) at points on the X axis. Remembering that r becomes x and $\cos\theta$ unity along the axis, we see by consulting Table I, page 89, that the desired series is

$$V = 2\pi i \left\{ 1 - \frac{r}{a}\cos\theta + \frac{1}{4}\frac{r^3}{a^3}(5\cos^3\theta - 3\cos\theta) + \cdots \right\}. \qquad (73\text{–}3)$$

Differentiating, we find for the components of magnetic intensity at P in the directions of increasing r and θ respectively,

$$H_r = -\frac{\partial V}{\partial r} = \frac{2\pi i}{a}\left\{ \cos\theta - \frac{3}{4}\frac{r^2}{a^2}(5\cos^3\theta - 3\cos\theta) + \cdots \right\},$$

$$(73\text{–}4)$$

$$H_\theta = -\frac{\partial V}{r\partial\theta} = \frac{2\pi i}{a}\left\{ -\sin\theta + \frac{3}{4}\frac{r^2}{a^2}(5\cos^2\theta - 1)\sin\theta + \cdots \right\}.$$

These expressions are useful in finding the field at a distance from the center of the circular current small compared to its radius. If, for instance, r is not greater than one quarter of a, the first two terms in the series specify the field with an error less than one percent.

Often it is desirable to obtain a field more uniform than that given by a circular current without having recourse to a long

solenoid. This may be effected by the use of two parallel circular currents of the same radius a placed a distance $2b$ apart as illustrated in Fig. 140. By a suitable separation of the circuits the field at a point O half-way between the two may be made more uniform than the field at any point due to either circuit alone. To find the optimum value of $2b$ we have from (70–3) for the field at a point P a distance x from O

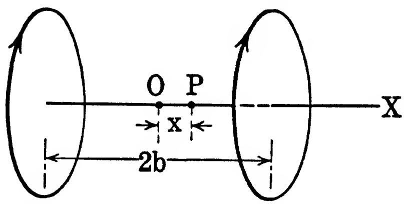


FIG. 140

$$H = 2\pi i a^2 \left\{ \frac{1}{[a^2 + (b + x)^2]^{3/2}} + \frac{1}{[a^2 + (b - x)^2]^{3/2}} \right\}. \quad (73\text{–}5)$$

Expanding by means of the binomial theorem in powers of x,

$$H = \frac{4\pi i a^2}{(a^2 + b^2)^{3/2}} \left\{ 1 + \frac{3(4b^2 - a^2)}{2(a^2 + b^2)^2} x^2 + \cdots \right\}. \quad (73\text{–}6$$

If $2b = a$ the coefficient of x^2 vanishes and the field is very uniform for some distance to either side of the mid-point O. Therefore the coils should be placed a distance apart equal to their radius. The magnetic intensity at O is then

$$H = \frac{32\pi i}{(5)^{3/2} a} = \frac{2.86\pi i}{a}. \quad (73\text{–}7)$$

Problem 73a. Find the field off the axis of a circular current at a distance r from the center of the circle greater than the radius a.

$$\text{Ans. } H_r = \frac{2\pi i}{a} \left\{ \frac{a^3}{r^3} \cos\theta - \frac{3}{4} \frac{a^5}{r^5} (5 \cos^3\theta - 3 \cos\theta) + \cdots \right\},$$

$$H_\theta = \frac{2\pi i}{a} \left\{ \frac{1}{2} \frac{a^3}{r^3} \sin\theta - \frac{9}{16} \frac{a^5}{r^5} (5 \cos^2\theta - 1) \sin\theta + \cdots \right\}.$$

Problem 73b. Show from the result of the previous problem that the field at a distance from a small circular current large compared to the radius of the circuit is the same as that due to a magnetic dipole of moment equal to the product of the current by the area of the circuit. Compare (70–9).

74. Galvanometers. — An instrument devised to measure the strength of a current is known as a *galvanometer*. We shall describe in this article some of the instruments used to measure currents which are based on the interaction between a current and a magnet. In some the circuit is fixed and the magnet, in the form of a compass needle, suffers a deflection when a current passes through the circuit; in others the magnet, in the form of a broken ring, is fixed, and the circuit, placed between the poles of the magnet, is deflected by the magnetic field of the latter. In either case the moving part must be subject to a control torque of restitution. Galvanometers may be classified as *absolute* or *sensitive*. The current passing through an instrument of the first type may be calculated in e.m.u. in terms of its dimensions, the strength of the field and the observed deflection. Instruments of the second type are designed to secure great sensitivity and are generally not absolute. Such an instrument must be calibrated by comparison with an absolute galvanometer connected in series with it.

Tangent Galvanometer. — This instrument consists of a vertical circular coil of n turns of radius a (Fig. 141) at the center of

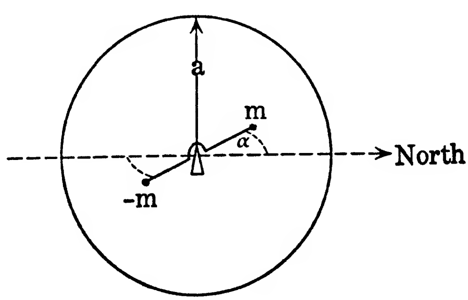


FIG. 141

which is suspended a compass needle free to turn in a horizontal plane about its center of mass. The length l of the needle must be small compared to a. The coil is oriented so that its plane is parallel to the earth's field, which supplies the control torque.

When a current i passes through the coil, the field at its center is

$$H = \frac{2\pi ni}{a}$$

along the axis of the coil, from (70–2). If H_e is the horizontal component of the earth's field, the compass needle assumes the direction of the resultant of the two perpendicular fields H and H_e. Consequently, if α is its deflection from the plane of the coil,

$$\tan\alpha = \frac{H}{H_e} = \frac{2\pi ni}{aH_e},$$

or

$$\begin{aligned} i &= \frac{aH_e}{2\pi n}\tan\alpha \\ &= K_i \tan\alpha, \end{aligned} \qquad (74\text{–}1)$$

where $K_i \equiv aH_e/2\pi n$ is known as the *galvanometer constant*. We have seen how to determine H_e in absolute units in article 47. Therefore i may be measured in absolute electromagnetic units by means of this instrument.

If the needle is not so short that the distance l between its poles is negligible compared to the radius a of the coil, the field at the poles differs slightly from the field at the center of the coil and we must make use of (73–4). The torque due to the current in the direction of decreasing θ is

$$\begin{aligned} -mlH_\theta &= \frac{2\pi nMi}{a}\sin\theta\left\{1 - \frac{3}{16}\frac{l^2}{a^2}(5\cos^2\theta - 1)\right\} \\ &= \frac{2\pi nMi}{a}\cos\alpha\left\{1 - \frac{3}{16}\frac{l^2}{a^2}(5\sin^2\alpha - 1)\right\}, \end{aligned}$$

since the deflection α of the needle from the plane of the coil is the complement of the angle θ of Fig. 139. The opposite torque due to the earth's field is $MH_e\sin\alpha$. Equating the two torques and solving for the current,

$$i = \frac{aH_e}{2\pi n}\tan\alpha\left\{1 + \frac{3}{16}\frac{l^2}{a^2}(5\sin^2\alpha - 1)\right\}. \qquad (74\text{–}2)$$

Since the poles of the needle are not always located at its ends it is difficult to determine the value of l needed for the correction term. The latter, however, may be made negligibly small by using the instrument so that the deflection is in the neighborhood of arc $\sin(1/\sqrt{5}) = 26°.6$.

Helmholtz has improved this type of galvanometer by placing the needle half-way between two parallel coils placed a distance apart equal to their common radius. As was shown in the last article the field is very uniform in the neighborhood of this point. Therefore no correction term is needed, the current being given by

$$i = \frac{(5)^{3/2} a H_e}{32\pi n} \tan \alpha \tag{74-3}$$

from (73–7), where n is the number of turns in each coil.

The current sensitivity of the tangent galvanometer is

$$\frac{d\alpha}{di} = \frac{2\pi n}{a H_e} \cos^2 \alpha = \frac{1}{K_i} \cos^2 \alpha \tag{74-4}$$

from (74–1). To make this great, a large number of turns, small radius and small control field H_e are needed, and the deflections should be small. Increasing the sensitivity by increasing n and decreasing a increases the magnitude of the necessary corrections. On the other hand, if the control field is made small the disturbing effect of stray fields, due to neighboring power circuits, is magnified.

A modification of the tangent galvanometer which secures increased sensitivity by decreasing the magnitude of the control torque consists of two fixed coils A and B (Fig. 142) in the same vertical plane wound so that the current traverses them in opposite senses. The needles (ns) at the centers of the coils are rigidly connected and free to rotate as a pair about a ver-

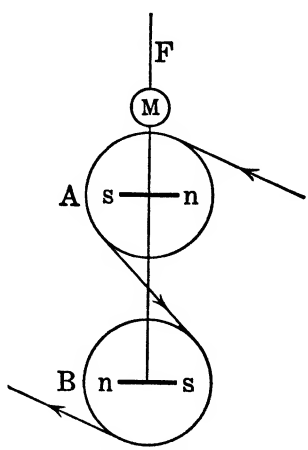


FIG. 142

tical axis. As they are of equal moments and oppositely oriented the earth's field exerts no torque on the movable part. Instead the control torque is supplied by the torsion fibre F. By using a fine fibre the torque of restitution can be made very small, thereby making the sensitivity very great. As the current traverses the two coils in opposite senses the torque exerted by it on each needle is in the same sense. Therefore the torque on the suspension due to the current is as great as in a single coil having a number of turns equal to the sum of those of the two coils. A mirror M is rigidly attached to the suspended system so that the deflection may be read with a telescope and scale in a manner to be described later. The whole apparatus is placed inside a cylindrical shell, or several coaxial shells, of soft iron so as to shield it from outside magnetic fields in accord with the theory of article 45. The type of instrument under discussion is known as an *astatic galvanometer*. As usually constructed it is not absolute.

If the needles are short compared to the radius a of each coil, the torque on the suspended system due to the current is

$$\frac{2\pi nMi}{a}\cos\alpha \tag{74-5}$$

for a deflection α, where M is the magnetic moment of either of the needles and n the sum of the number of turns in the two coils. Equating this to the torque of restitution $k\alpha$ of the twisted fibre,

$$i = \frac{ak}{2\pi nM\cos\alpha}\alpha. \tag{74-6}$$

If the deflection is small $\cos\alpha$ remains sensibly unity, and

$$\begin{aligned} i &= \frac{ak}{2\pi nM}\alpha \\ &= K_i\alpha, \end{aligned} \tag{74-7}$$

where $K_i \equiv ak/2\pi nM$ is the galvanometer constant for this type of instrument.

D'Arsonval Galvanometer. — The moving needle type of galvanometer is easily disturbed by stray fields. This difficulty is

obviated in the moving coil instrument depicted in plan in the upper part of Fig. 143 and in vertical section in the lower part of the figure. The flat rectangular coil $BCDE$ of n turns surrounding a fixed cylinder of soft iron, is suspended by the torsion fibre F between the poles NS of a strong magnet so as to lie in a plane parallel to the lines of force when no current is flowing. The current i enters the coil through the upper suspension and leaves below.

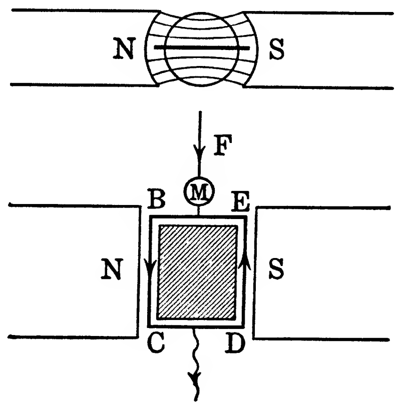


FIG. 143

If l is the length $DE = BC$ the force on DE due to the field H of the magnet is iHl in the backward direction from (69–9). An equal forward force is experienced by BC. If h is the width $EB = CD$ the torque due to this couple is $iHlh$. This is the total torque on one turn, as the currents in the horizontal arms are subject to no turning moment. Multiplying by the number of turns the torque on the entire coil is seen to be $niHA$, where A is the area lh of the rectangle.

The soft iron core which the coil surrounds concentrates the lines of induction in its interior as illustrated in Fig. 143. Therefore the field is nearly radial for small deflections of the coil and the torque which it exerts on the current remains unchanged. For larger deflections, however, the vertical arms of the coil come into a portion of the field which is more nearly transverse and the lever arm of the couple is therefore approximately $h \cos \alpha$, where α is the deflection, instead of h. Consequently the torque becomes $niHA \cos \alpha$.

The torque of restitution exerted by the torsion fibre is proportional to the deflection and may be written $k\alpha$, where k is a constant. Equating this to the torque on the coil, we have for

small deflections

$$i = \frac{k}{nAH}\alpha$$
$$= K_i\alpha, \tag{74-8}$$

where $K_i \equiv k/nAH$ is the galvanometer constant, and for large deflections,

$$i = \frac{k}{nAH \cos \alpha}\alpha. \tag{74-9}$$

In Fig. 144 the current is plotted against the deflection, the lower broken line corresponding to (74–8) and the upper one to (74–9). The actual curve observed, shown by the full line, lies between these, following the straight line very closely near the origin.

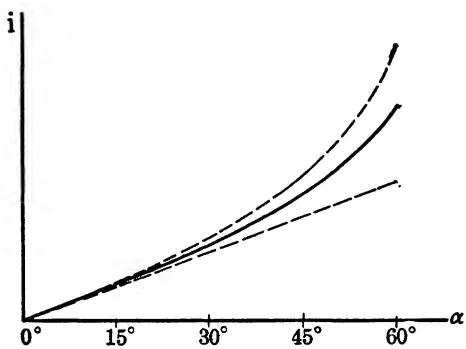


FIG. 144

The current sensitivity of the D'Arsonval galvanometer for small deflections is

$$\frac{d\alpha}{di} = \frac{nAH}{k} = \frac{1}{K_i}. \tag{74-10}$$

To obtain high sensitivity the coil should have many turns of fine wire, the field H should be large, and the stiffness of the suspension small. For this reason the coil is suspended by a fine wire usually of phosphor-bronze. The sensitivity may be made great enough to measure currents of $(10)^{-11}$ ampere. As the field of the fixed magnet is large, stray fields produce relatively unimportant disturbances. The instrument, as generally constructed, is not absolute and must be calibrated by comparison with an absolute galvanometer.

Sensitive galvanometers of this type are usually provided with a mirror M (Fig. 143) which turns with the coil. A telescope T

and scale S (Fig. 145) are placed in front of the mirror. A deflection α of the mirror brings on the cross-hair of the telescope a point P on the scale at an angular distance 2α from O. We generally express K_i in amperes per scale division to avoid calculating α.

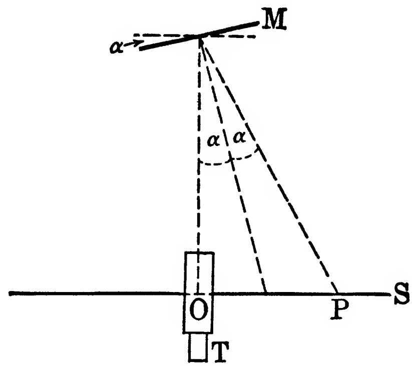


FIG. 145

The ammeters and voltmeters used for direct current measurements are merely D'Arsonval galvanometers provided with a pointer instead of mirror and telescope, the torque of restitution being supplied by a fine spiral spring instead of a torsion fibre. Such an instrument is shown with the case removed in Fig. 146.

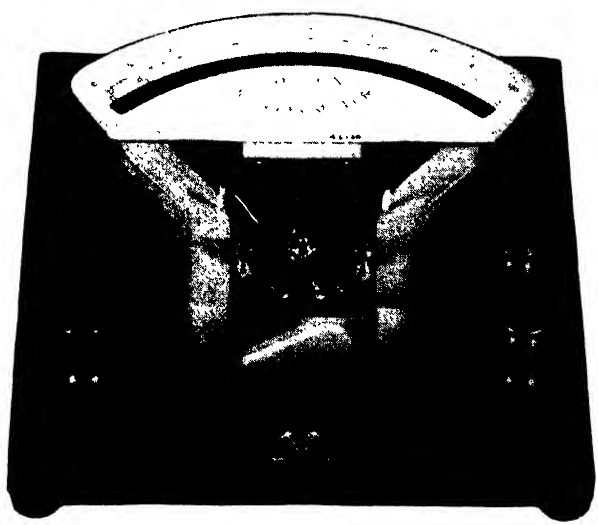


FIG. 146

String Galvanometer. — The string galvanometer, invented by Einthoven, has the property of being very sensitive and at the

same time rapid in taking up its final deflection. It is very useful in measuring small currents which are changing in magnitude.

This instrument consists of a single conducting fibre AB (Fig. 147), usually silvered quartz, placed under tension between the poles NS of a strong magnet. A hole ED is drilled through one of the pole pieces so that the fibre can be viewed by means of a microscope. When a current passes along the fibre in the direction of the arrow a force acts on it tending to deflect it away from the reader. The deflection is measured by means of a movable cross-hair in the image plane of the microscope. Like the moving coil galvanometer, this instrument is comparatively free from disturbances due to stray fields and is not absolute. It may be made sensitive enough to detect currents of $(10)^{-12}$ ampere.

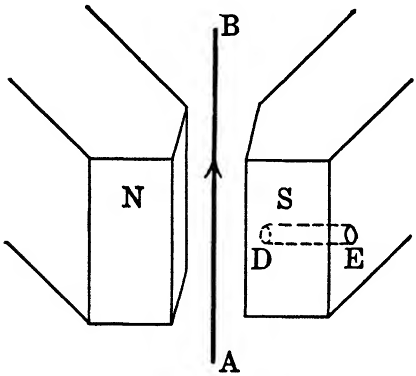


Fig. 147

Problem 74a. Show that if the coil of the tangent galvanometer is rotated about a vertical axis so as to lie in the same vertical plane as the needle when the deflection is read (sine galvanometer),

$$i = \frac{aH_e}{2\pi n} \sin \alpha.$$

Compare the sensitivity of this galvanometer with that of the tangent instrument.

75. Ampère's Theory of Magnetism. — We have seen that a small current circuit produces a magnetic field which is indistinguishable, at distances large compared to the dimensions of the circuit, from that of a magnetic dipole. In fact, it has been shown that the field outside a magnetic shell of constant strength is identical with that of a current equal to the strength of the

shell, flowing around a circuit coinciding with the periphery of the shell. These considerations led Ampère to suggest that the magnetic moments exhibited by the molecules of paramagnetic and ferromagnetic substances are due not to the presence of magnetic dipoles but to the existence of intra-molecular currents. As the magnetic properties of such Ampèrian circuits are permanent, the currents involved must be supposed to flow without resistance. Ampère's theory, which is universally accepted today, makes unnecessary the supposition of the existence of a magnetic entity, all magnetic phenomena being attributed to moving electrical charges. Our present knowledge of the structure of atoms leads us to believe that the Ampèrian circuits existing in the molecules of magnetic media consist of rings of electrons revolving about the nuclei of the atoms, or possibly of electrons spinning, like the earth, about a diameter as axis.

While it is a matter of indifference in describing the field outside a magnet whether we attribute magnetic properties to actual dipoles or to intra-molecular currents, the two points of view lead to different conclusions regarding the field in the interior of the substance. Therefore we must investigate the field inside a magnetic medium from the point of view of Ampère's theory. As it is only the mean field which we can detect experimentally, our problem is to discover how the mean field inside a medium whose molecules contain Ampèrian circuits differs from that inside a medium whose molecules contain equivalent magnetic shells.

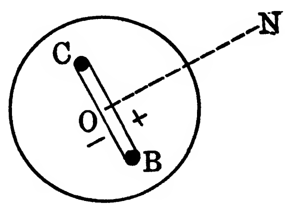


FIG. 148

Consider a single atom (Fig. 148) containing a plane Ampèrian circuit shown in cross-section by BC, the current i entering the plane of the paper at B and emerging at C. If we replace the current by an equivalent magnetic shell whose periphery coincides with the circuit, the field is altered only in the interior of the shell. Let the thickness of the shell be Δl. Then, as the strength of the shell is i, the intensity of

magnetization in its interior is

$$I_0 = \frac{i}{\Delta l}.$$

Now the shell has a magnetic charge I_0 per unit area on the positive face and $-I_0$ per unit area on the negative face from (42–4). It is therefore the magnetic analog of a parallel plate condenser, and as was shown in problem 39*a* the field in the interior is $4\pi I_0$ in the direction opposite to the normal ON. If A is the area of the shell and τ_0 the volume of the atom, the contribution of the field inside the shell to the mean field averaged over the volume of the atom is

$$H_0 = \frac{1}{\tau_0}(4\pi I_0)(A\Delta l) = \frac{4\pi i A}{\tau_0}$$

in the direction opposite to that of the positive normal ON. As iA is the magnetic moment M of the shell and τ_0 the reciprocal of the number n of atoms per unit volume, we can write this in vector form as follows:

$$\mathbf{H}_0 = -4\pi n\mathbf{M},$$

the minus sign indicating that $\mathbf{H}_0$ has the opposite direction to $\mathbf{M}$.

So far we have averaged over a single atom only. If now we average over a volume τ containing a large number of atoms,

$$\mathbf{H}_0 = -4\pi n\overline{\mathbf{M}},$$

the bar indicating the mean value of the quantity over which it appears. But $n\overline{\mathbf{M}}$ is just the mean magnetic moment per unit volume, that is, the intensity of magnetization $\mathbf{I}$ of the medium. So the contribution to the mean magnetic intensity of the fields in the interiors of the magnetic shells by which we have replaced the Ampèrian circuits actually existing inside the atoms is

$$\mathbf{H}_0 = -4\pi\mathbf{I}. \tag{75–1}$$

We wish to compare (75–1) with the contribution to the mean field provided by the portion of the field of the Ampèrian circuits which occupies the space previously filled by the equiva-

lent shells. As the field due to a circuit is continuous all the way around we can, however, make the latter contribution vanishingly small by taking Δl small enough. Therefore we conclude that the mean magnetic intensity inside a medium composed of Ampèrian circuits exceeds that inside an exactly similar medium composed of equivalent magnetic shells by $4\pi\mathbf{I}$. If $\mathbf{H}_i$ is the magnetic intensity in the first and $\mathbf{H}_m$ that in the second,

$$\mathbf{H}_i = \mathbf{H}_m + 4\pi\mathbf{I}. \tag{75-2}$$

Now $\mathbf{H}_m + 4\pi\mathbf{I}$ is the quantity that we have designated by **B** and named the *magnetic induction*. Therefore we conclude that on Ampère's theory of magnetism the mean magnetic intensity inside a magnetic medium, as well as outside, is the magnetic induction **B**. The quantity which we have designated by **H** in the discussion of magnetic media in Chapter IV is not the magnetic intensity actually existing inside a magnetic medium at all, but the fictitious field which would exist if each Ampèrian circuit were replaced by an equivalent magnetic shell. Outside magnetic media **B** and **H** are the same and either represents the true magnetic intensity.

The conclusions reached in this article enable us to generalize the expressions found in article 69 for the force exerted on a current element by an external magnetic field so as to apply to currents flowing through magnetic media. Since the mean magnetic intensity inside such a medium is **B**, equation (69–8) becomes

$$d\mathbf{F} = \mathbf{i} \times \mathbf{B}dl, \tag{75-3}$$

and the expression (69–10) for the force per unit volume,

$$\mathbf{F}_\tau = \mathbf{j} \times \mathbf{B}. \tag{75-4}$$

Problem 75a. A steady current i flows along a straight wire of circular cross-section composed of a paramagnetic substance of permeability μ. If a is the radius of the wire, find B both inside and outside the wire. As the surface of the wire is crossed do B and H change continuously or abruptly? Draw graphs of both B and H plotted against the distance R from the axis of the wire.

Ans. $\dfrac{2\mu i}{a^2}R,\quad \dfrac{2i}{R}$.

CHAPTER VIII

MOTION OF IONS IN ELECTRIC AND MAGNETIC FIELDS

76. Equation of Motion. — We shall use the term *ion* in this chapter to designate any free charged particle. It may be an atom, molecule or aggregation of atoms or molecules which has acquired a positive charge through the loss of one or more electrons or a negative charge through the attachment of extra electrons, or it may be a free electron, such as a β ray emitted by a radioactive atom, or a free proton. We shall limit our discussion to ions which have velocities relative to the observer small compared to the velocity of light, and for the present we shall suppose that the ions under consideration are so far apart and moving through a region where the gas pressure is so low that we can neglect (*a*) collisions of ions with other ions or with neutral molecules, (*b*) the forces exerted by the ions on one another.

We shall find it convenient to write our equations entirely in electromagnetic units. We have defined the electromagnetic unit of charge in article 69. The electromagnetic unit of electric intensity is that existing when the electrical force on an electromagnetic unit of charge is one dyne. Therefore the electrical force on an ion charged with e electromagnetic units of charge moving through an electric field where the electric intensity is $\mathbf{E}$ electromagnetic units is $e\mathbf{E}$. The magnetic force on a moving charge is given by (69–11). For an ion of charge e this force is $e\mathbf{v} \times \mathbf{H}$, where $\mathbf{v}$ represents the velocity of the ion relative to the observer.

Adding the electric and magnetic forces together the total force on an ion with charge e moving with velocity $\mathbf{v}$ is

$$\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{H}), \tag{76–1}$$

all the quantities on the right being expressed in electromagnetic units. The force due to the magnetic field exists only when the ion is moving relative to the observer. It is directed at right angles to both $\mathbf{v}$ and $\mathbf{H}$ in the direction of advance of a right-handed screw rotated from the first to the second of these vectors. Being at right angles to $\mathbf{v}$ the force due to the magnetic field never does any work on the moving ion. As the force due to the magnetic field is proportional to the sine of the angle between $\mathbf{v}$ and $\mathbf{H}$ it is maximum when the motion is at right angles to the lines of magnetic force and zero when the motion is along the lines of force. If the ion should be moving in a magnetic medium, as in the case of a current passing through a magnet, $\mathbf{H}$ must be replaced by $\mathbf{B}$ as noted in article 75, giving for the force on the ion,

$$\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \tag{76–2}$$

If, then, we can treat each ion as an isolated particle not subject to collisions with neighboring ions or neutral particles, the equation of motion of an ion of mass m moving through an electric field $\mathbf{E}$ and a magnetic field $\mathbf{B}$ is

$$m\mathbf{f} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \tag{76–3}$$

the vector $\mathbf{f}$ representing the acceleration of the ion.

This equation and all others in this chapter are expressed in electromagnetic units, except those in article 80, which are in electrostatic units.

77. Uniform Fields. — We shall consider three cases under this heading: (*a*) a uniform electric field alone, (*b*) a uniform magnetic field alone, (*c*) combined uniform electric and magnetic fields.

Uniform Electric Field. — The equation of motion of an ion in a uniform electric field $\mathbf{E}$ is

$$\mathbf{f} = \frac{e}{m}\mathbf{E}. \tag{77–1}$$

The ion has a constant acceleration in the direction of the lines of force. If it starts from rest it acquires a speed ft in a

time t and traverses a distance $\frac{1}{2}ft^2$. If it is projected with an initial velocity which is not parallel to the lines of force it describes a parabola of the same type as that of a projectile moving without resistance through the earth's gravitational field.

Uniform Magnetic Field. — The equation of motion of an ion moving through a uniform magnetic field in the absence of an electric field is

$$\mathbf{f} = \frac{e}{m}\mathbf{v} \times \mathbf{H}. \tag{77-2}$$

As the acceleration is always perpendicular to the velocity the speed of the particle remains constant throughout its motion. Moreover, as the acceleration is perpendicular to $\mathbf{H}$, the component of the velocity along the lines of force does not change with the time. Therefore it is only the motion at right angles to the field which requires consideration, and we can limit our discussion for the moment to the special case where $\mathbf{v}$ is perpendicular to $\mathbf{H}$. Then the magnitude of $\mathbf{v} \times \mathbf{H}$ is vH and as $\mathbf{f}$ is constant in magnitude and at right angles to $\mathbf{v}$ the ion describes a circle (Fig. 149). If ρ is the radius of this circle the normal acceleration f has the value v^2/ρ and the equation of motion becomes

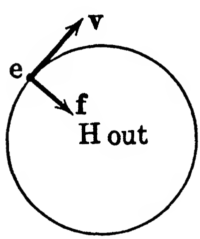


FIG. 149

$$\frac{v^2}{\rho} = \frac{e}{m}vH.$$

Consequently the radius of the circle is

$$\rho = \frac{mv}{eH}, \tag{77-3}$$

proportional to the initial velocity of the ion. The angular velocity with which the circulaı path is described is

$$\omega = \frac{v}{\rho} = \frac{eH}{m},$$

independent of the initial velocity. If $\mathbf{H}$ is out from the paper in Fig. 149 the circular path is described in the clockwise sense by a positively charged particle. Hence $\boldsymbol{\omega}$ is parallel but opposite to $\mathbf{H}$, and, writing the last equation in vector form,

$$\boldsymbol{\omega} = -\frac{e}{m}\mathbf{H}. \tag{77-4}$$

A negatively charged ion, then, describes a circular path in the counter-clockwise sense.

If, now, the initial velocity of the ion is not perpendicular to the lines of force, equation (77–3) holds if we understand by v the normal component of the velocity. In addition to motion about the lines of force with angular velocity $\boldsymbol{\omega}$ specified by (77–4) in a circle of radius ρ given by (77–3) the ion has a uniform translation parallel to the lines of force with a speed equal to the component of the initial velocity parallel to the field. Therefore the path is a helix with its axis in the direction of $\mathbf{H}$.

By allowing a stream of ions to pass through a uniform magnetic field the velocities of the individual particles can be calculated by (77–3) from the observed deflections. This method is used in obtaining the *magnetic spectrum* of β rays (electrons) from radio-active sources. The rays, proceeding from the source S (Fig. 150) pass through the slit L into a magnetic field at right angles to the plane of the paper, and, after completing a semicircle, impinge on the photographic plate P. On account of the finite angular width of the beam, different electrons describe different circular paths, but all those of the same velocity cross at approximately the same point after completing a half revolution. This focussing action results in a single line at P corresponding to a group of electrons of a definite initial velocity, another group of greater velocity coming to a focus at Q and so on.

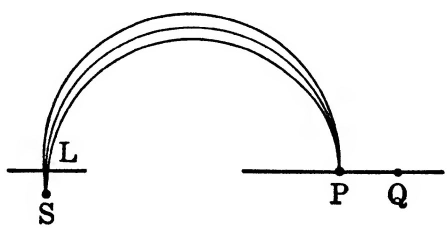


FIG. 150

Combined Electric and Magnetic Fields. — We shall consider now the motion of an ion in a region where it is acted on by both a uniform electric field and a uniform magnetic field. As no force is exerted by the magnetic field in the direction of the lines of force, the motion parallel to $\mathbf{H}$ is uniformly accelerated motion produced by the component of $\mathbf{E}$ parallel to the magnetic lines of force. The only part of the motion which requires examination, then, is that in the plane perpendicular to the magnetic intensity. Consequently we may confine our attention to the special case where the two fields are perpendicular and the initial velocity is at right angles to $\mathbf{H}$.

Let us orient the right-handed axes xyz fixed in the observer's inertial system so that y is parallel to $\mathbf{E}$ and z to $\mathbf{H}$. Consider, now, a second set of parallel axes XYZ which are moving relative to the observer's system in the x direction with a constant velocity

$$u = \frac{E}{H},$$

that is, in vector notation,

$$\mathbf{u} = \frac{\mathbf{E} \times \mathbf{H}}{H^2}. \tag{77-5}$$

Denoting by $\mathbf{v}$ the velocity of the ion relative to xyz and by $\mathbf{V}$ its velocity relative to the moving axes XYZ,

$$\mathbf{v} = \mathbf{u} + \mathbf{V}. \tag{77-6}$$

Now the equation of motion of the ion relative to the observer is

$$\frac{d\mathbf{v}}{dt} = \frac{e}{m}(\mathbf{E} + \mathbf{v} \times \mathbf{H}). \tag{77-7}$$

From (77–6) we have for the acceleration

$$\frac{d\mathbf{v}}{dt} = \frac{d\mathbf{V}}{dt}. \tag{77-8}$$

Substituting (77–6) and (77–8) in (77–7) we find for the equation of motion relative to the moving axes

$$\frac{d\mathbf{V}}{dt} = \frac{e}{m}\{\mathbf{E} + \mathbf{u} \times \mathbf{H} + \mathbf{V} \times \mathbf{H}\}. \tag{77-9}$$

Now

$$\mathbf{u} \times \mathbf{H} = \frac{(\mathbf{E} \times \mathbf{H}) \times \mathbf{H}}{H^2} = -\frac{\mathbf{E}H^2}{H^2} = -\mathbf{E},$$

as is clear from Fig. 151. Therefore (77–9) becomes

$$\frac{d\mathbf{V}}{dt} = \frac{e}{m}\mathbf{V} \times \mathbf{H}. \qquad (77\text{–}10)$$

By passing to a set of axes moving relative to the observer with the velocity **u** at right angles to the two fields we have eliminated **E**. An observer moving with these axes would detect no electric field. He would assert that the ion under consideration was moving under the influence of a magnetic field alone. This example illustrates the statement made in article 6 that the description of an electromagnetic field depends as much upon the state of motion of the observer as upon the distribution of charges, poles and currents.

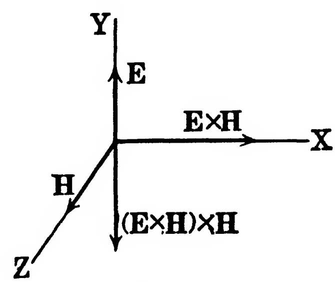


FIG. 151

As equation (77–10) for the motion of the ion is identical with equation (77–2) for the case of zero electric field, we know at once that the ion under consideration describes a circle of radius

$$\rho = \frac{mV}{eH} \qquad (77\text{–}11)$$

in the XY plane with angular velocity

$$\omega = -\frac{e}{m}\mathbf{H} \qquad (77\text{–}12)$$

relative to the moving axes XYZ, the speed V with respect to the moving axes remaining constant during the motion. Relative to the axes xyz fixed in the observer's inertial system the center of this circle moves in the x direction with a uniform velocity of translation u. Therefore the path of the ion is the cycloid de-

scribed by a point at a distance ρ from the center of a circle rolling on the x axis. As the velocity of the center of the rolling circle is u its radius is

$$a = \frac{u}{\omega} = \frac{m}{e}\frac{E}{H^2}, \tag{77-13}$$

independent of the initial velocity of the ion. The ratio of ρ to a is

$$\frac{\rho}{a} = \frac{V}{u}. \tag{77-14}$$

According as V is greater than, equal to, or less than u the generating point lies outside, on the circumference of, or inside the rolling circle and the cycloid is curtate, common, or prolate. Typical paths are shown in Fig. 152 for the case of a positive

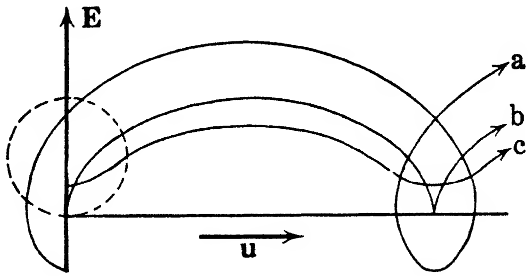


FIG. 152

ion, **H** being directed toward the reader. The rolling circle is indicated by a broken line, (a) representing a curtate, (b) a common, and (c) a prolate cycloid. The drift **u** resembles the precession of a rotating top in that it is at right angles to the impressed fields **E** and **H**.

Since **u** depends only on **E** and **H**, it is in the same sense for a negative ion as for a positive ion and is independent of the magnitude of charge or mass. The paths for negative ions are obtained by rotating those in Fig. 152 through the angle π about the horizontal axis.

If $\mathbf{v}_0$ is the initial velocity of an ion relative to the observer's inertial system and θ the angle which it makes with the x axis,

$$V^2 = v_0^2 - 2uv_0 \cos\theta + u^2.$$

from (77–6), or

$$V^2 - u^2 = v_0(v_0 - 2u \cos \theta). \tag{77–15}$$

We can classify all possible paths in terms of the initial velocity as follows:

(1) $V = 0$. This case exists only when $\theta = 0$ and $v_0 = u$. For if we make V zero in (77–15) and solve for v_0,

$$v_0 = u(\cos \theta \pm \sqrt{\cos^2 \theta - 1}),$$

which is real and positive only for $\cos \theta$ equal to unity, giving v_0 equal to u. The path is a straight line at right angles to **E** and **H**, the force due to the electric field being exactly balanced by the opposite force due to the magnetic field.

(2) $V \neq 0$. We have three subcases, namely;

(*a*) $v_0 = 0$. The path is a common cycloid, the ion starting at the cusp.

(*b*) $0 < v_0 \leq 2u$. The path is a prolate cycloid, common cycloid, or curtate cycloid according as the angle θ is less than, equal to, or greater than that defined by $\cos \theta = v_0/2u$, for these are the conditions that V should be less than, equal to, or greater than u as is evident from (77–15).

(*c*) $2u < v_0$. The path is a curtate cycloid for all values of θ.

In the preceding discussion we have supposed the electric and magnetic fields to be at right angles and the initial velocity to lie in the plane perpendicular to the magnetic field. If these conditions are not satisfied we must understand by **E** in our formulas the component of the electric intensity at right angles to **H** and by the various velocities involved the component velocities in the plane perpendicular to **H**. Then, in addition to the motion depicted in Fig. 152, the ion has a uniform acceleration at right angles to the plane of the figure due to the component of the electric intensity parallel to the magnetic field.

By deflecting a stream of ions by crossed electric and magnetic fields at right angles to the initial velocity of the stream both the

velocity and the ratio of charge to mass of the ions may be measured. The electric field may be produced by the parallel plate condenser AB (Fig. 153) between the plates of which the stream of ions CD passes, the magnetic field (indicated by dots) at right angles to the plane of the figure being produced by a solenoid.

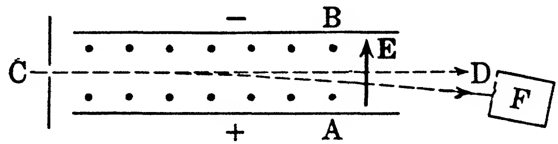


FIG. 153

First the potential difference of the condenser is adjusted until the stream suffers no deflection. Under these conditions the upward force due to the electric field is just balanced by the downward force due to the magnetic field, and the velocity v of the ions is equal to the drift velocity u, that is,

$$v = u = \frac{E}{H}.$$

If, now, the electric field is suppressed, the ions describe the circular arc CF of radius

$$\rho = \frac{mv}{eH}.$$

The radius ρ is obtained at once from the observed deflection, and as v is known, e/m can be calculated by means of the formula

$$\frac{e}{m} = \frac{v}{\rho H}, \tag{77-16}$$

obtained by solving the preceding equation for the ratio of charge to mass.

The deflection may be observed by placing at F a metal cylinder with a narrow opening, known as a *Faraday cylinder*, which is connected to an electrometer. The strength of the magnetic field is varied until the electrometer shows a maximum

rate of deflection, indicating that all the ions in the stream are entering the chamber. Another method of measuring the deflection is by means of a photographic plate placed at D at right angles to the stream.

By employing this method J. J. Thomson was able to measure, in 1897, the ratio e/m of electrons coming from the cathode of an evacuated tube through which a discharge is passing. The most recent value of the ratio of charge to mass of the electron obtained by the deflection method is

$$\frac{e}{m} = -1.769(10)^7 \text{ e.m.u./gm} = -5.30(10)^{17} \text{ e.s.u./gm.} \quad (77\text{–}17)$$

Thomson's experiments also showed that the velocity of cathode rays is of the order of one-tenth that of light.

Sometimes, as in the *positive ray spectrograph*, it is convenient to have the electric and magnetic fields parallel. Suppose both

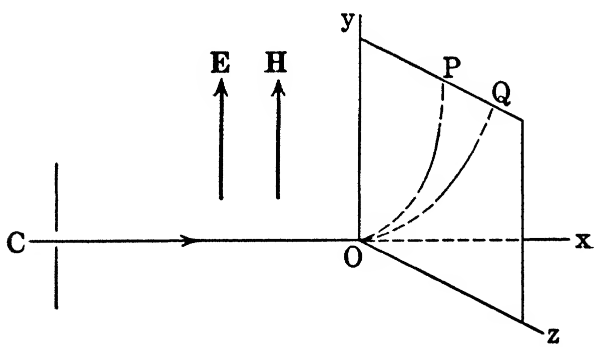


FIG. 154

fields have the direction of the y axis (Fig. 154), the ions entering the field at C along the x axis. Assuming both deflections to be small, the deflection due to the electric field at the end of the time t is

$$y = \frac{1}{2}\frac{eE}{m}t^2,$$

and that due to the circular path (77–3) imposed by the mag-

netic field is

$$z = \frac{1}{2}\frac{eHv}{m}t^2 = \frac{1}{2}\frac{eHs}{m}t,$$

where $s = vt$ is the length of the path through the fields. Eliminating t we get the parabola

$$y = \frac{2E}{H^2 s^2}\frac{m}{e}z^2.$$

All ions having the same value of e/m strike a photographic plate lying in the yz plane somewhere along the arc OP of this parabola, whatever their velocities may be. Ions with a greater value of e/m give rise to another parabolic arc, such as OQ. If the particles under consideration are singly ionized, the charge e is the same for all. We have here, then, a method of differentiating the masses of ions which does not require a knowledge of their velocities.

By passing charged atoms first through an electric field and then through a magnetic field at right angles in an instrument known as the *mass spectrograph*, Aston has shown that many of the elements are mixtures of atoms of two or more species, each of which has a characteristic atomic weight. Such species are known as *isotopes*, since they occupy the same place in the periodic table. Thus chlorine consists of two isotopes, one with atomic weight 35 and the other with atomic weight 37. On account of the greater prevalence of the former chemical methods of determining the atomic weight give 35.46.

Problem 77a. If α is the angle which the velocity $\mathbf{v}$ of an ion projected in a uniform magnetic field makes with the lines of force show that the pitch of the helix described is $\frac{2\pi m}{eH} v \cos \alpha$.

Problem 77b. Under the influence of ultra-violet light electrons are emitted normally with negligible velocities from one plate of a parallel plate condenser placed in a magnetic field H parallel to the plates. Find the relation between the potential drop V across the condenser, the separation D of the plates, and the field H so that the electrons may just get across.

$$\text{Ans. } V = \frac{1}{2}\frac{e}{m}D^2H^2.$$

Problem 77c. What is the drift velocity of ions moving in crossed gravitational ($m\mathbf{g}$) and magnetic ($\mathbf{H}$) fields? Do ions of both signs progress in the same sense?

$$\text{Ans. } \mathbf{u} = \frac{m}{e}\frac{\mathbf{g}\times\mathbf{H}}{H^2}, \text{ Opposite senses.}$$

Problem 77d. Find the integrated equations of motion of an ion moving in crossed electric and magnetic fields, $\mathbf{E}$ being parallel to the y axis and $\mathbf{H}$ to the z axis and the initial velocity being denoted by $\mathbf{v}_0$.

$$\text{Ans. } x = \frac{m}{e}\left\{\frac{v_{0y}}{H}\left(1 - \cos\frac{eH}{m}t\right) + \left(\frac{v_{0x}}{H} - \frac{E}{H^2}\right)\sin\frac{eH}{m}t\right\} + \frac{E}{H}t,$$

$$y = \frac{m}{e}\left\{-\left(\frac{v_{0x}}{H} - \frac{E}{H^2}\right)\left(1 - \cos\frac{eH}{m}t\right) + \frac{v_{0y}}{H}\sin\frac{eH}{m}t\right\},$$

$$z = v_{0z}t.$$

78. Charge and Mass of the Electron. — In the last article we have seen how the ratio of the charge to the mass of the electron may be measured by means of the deflection of a stream of cathode rays in crossed electric and magnetic fields. We shall now describe the method used by Millikan for measuring the charge of the electron. From this measurement and that of the ratio of charge to mass we can calculate the mass of the electron, as well as certain other important physical constants.

Suppose oil drops of radii between $(10)^{-5}$ and $(10)^{-4}$ cm to be sprayed into the space between the horizontal plates (Fig. 155) of a parallel plate condenser. If the air between the plates is ionized by a flash of X-rays or by other means, occasionally one or more ions will attach themselves to an oil drop, the motion of which can be observed through a microscope provided with cross-hairs. By charging the condenser to a potential V, the oil drop may be subjected to an upward electrical force eE, where e is the charge on the drop and $E = V/d$ is the electric intensity.

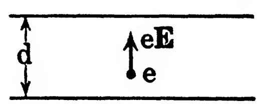


Fig. 155

For simplicity we shall suppose that only a single electron is attached to the drop. Then e is the charge on the electron. It is a simple matter to differentiate experimentally the cases where

two or three electrons have become fastened to the drop from the case where only a single electron is involved.

The forces acting on the drop are the downward force of gravity, the upward force due to the buoyancy of the air, and the upward force due to the electric field. Let a be the radius of the drop, which must assume a spherical form on account of surface tension, σ the density of the oil of which it is composed, and ρ the density of the air. Then the three forces just enumerated are $\frac{4}{3}\pi a^3\sigma g$, $\frac{4}{3}\pi a^3\rho g$, and eE respectively. So, if the potential is adjusted until the drop is in equilibrium,

$$\tfrac{4}{3}\pi a^3(\sigma - \rho)g = eE. \tag{78-1}$$

While σ, ρ and E are easily measured we cannot calculate e until we determine the radius a of the drop. To do this we remove the electric field and observe the speed v with which the drop falls. A deduction from hydrodynamics, known as *Stokes' law*, shows that v is related to a and the viscosity η of the air through which the drop is falling by the equation

$$a = \sqrt{\frac{9}{2}\frac{\eta}{\sigma - \rho}\frac{v}{g}}. \tag{78-2}$$

Eliminating a from (78-1) by means of this relation and denoting the value of e so obtained by e_1,

$$e_1 = \frac{4\pi}{3}\left(\frac{9\eta}{2}\right)^{3/2}\left\{\frac{1}{g(\sigma - \rho)}\right\}^{1/2}\frac{v^{3/2}}{E}. \tag{78-3}$$

If the charge on the electron is computed from this formula it is found that the value obtained depends upon the pressure of the air and the radius of the drop. The trouble is that the drops are so close to molecular dimensions that the hydrodynamical assumption underlying Stokes' law that the fluid through which the drop is falling has a continuous structure is no longer altogether valid. Consequently (78-3) can be considered as correct only in the limiting case of large drops or high pressures of the gas. As the pressure of a gas is inversely proportional to the mean free path l of the molecules composing it,

our formula may be expected to hold accurately only for small values of the ratio l/a. Now (78–3) shows that v is proportional to $e_1^{2/3}$. As it is the expression for v given by Stokes' law which requires correction, we can investigate the nature of the correction empirically by plotting the values of $e_1^{2/3}$ obtained from (78–3) for drops of different sizes at different gas pressures against l/a. Doing so, the straight line indicated in Fig. 156 is obtained. Extending this line to the axis of ordinates $e^{2/3}$ is found.

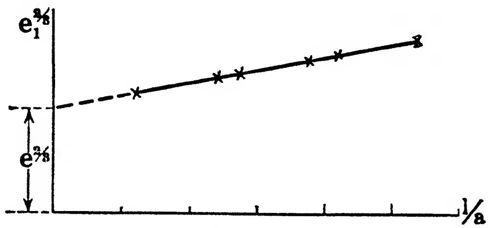


FIG. 156

In this way Millikan finds for the charge on the electron

$$e = -1.591(10)^{-20} \text{ e.m.u.} = -4.770(10)^{-10} \text{ e.s.u.} \qquad (78\text{–}4)$$

Combining this with (77–17) we compute the mass of the electron to be

$$m = 8.994(10)^{-28} \text{ gm.} \qquad (78\text{–}5)$$

The deflection method of the previous article shows that the ratio of charge to mass is 1847 times as great for a hydrogen nucleus or proton as for an electron. Therefore the mass of the proton (or hydrogen atom) is

$$m = 1.661(10)^{-24} \text{ gm.} \qquad (78\text{–}6)$$

As noted in article 57 Faraday's constant F representing the charge carried in electrolysis by one gram-equivalent is 96,490 coulombs or 9649 e.m.u. of charge. Therefore Loschmidt's constant, that is, the number of atoms per gram-atom or of molecules per gram-molecule is

$$L = \frac{9649}{1.591(10)^{-20}} = 6.064(10)^{23} \text{ molecule/gm-molecule.} \qquad (78\text{–}7)$$

Since one gram-molecule of an ideal gas is found to occupy $2.241(10)^4$ cm^3 at 0° C and one atmosphere pressure, the number of molecules per cubic centimeter under standard conditions (Avogadro's constant) is

$$A = \frac{6.064(10)^{23}}{2.241(10)^4} = 2.705(10)^{19} \text{ molecule/cm}^3. \qquad (78\text{–}8)$$

The equation of state of an ideal gas composed of N molecules is

$$pV = kNT,$$

where the Boltzmann constant k is one that appears in many important physical relations. As (78–8) gives us

$$A = \frac{N}{V} = 2.705(10)^{19} \text{ molecule/cm}^3,$$

for the temperature $T = 273°.1$ K and the pressure $p = 1$ atmos. $= 1.0132(10)^6$ dyne/cm^2, we can calculate k, getting

$$k = \frac{1.0132(10)^6}{2.705(10)^{19} \times 273.1} = 1.371(10)^{-16} \text{ erg/deg C.} \qquad (78\text{–}9)$$

If M represents the number of gram-molecules of an ideal gas in a volume V it follows from Avogadro's law that the gas constant R in the equation $pV = RMT$ is a universal constant. Its value is

$$R = kL = 8.314(10)^7 \text{ erg/deg C gm-molecule.} \qquad (78\text{–}10)$$

79. The Hall Effect. — If the terminals of a galvanometer G (Fig. 157) are attached to two opposite points A and B of a thin metal strip MN carrying a current i, the galvanometer will remain undeflected, showing that the equipotential line AB is perpendicular to the current. If, however, the strip is placed in a magnetic field H at right angles to its plane, a deflection is observed which does not vanish until the lead at B is displaced to a point C beyond B. The equipotential lines, therefore, are

rotated from the direction of AB to that of AC by the magnetic field. This effect was discovered by E. H. Hall in 1879.

As the free paths of the electrons responsible for metallic conduction are probably short, these electrons describe only small portions of the cycloidal arcs depicted in Fig. 152 when subject to crossed electric and magnetic fields. Suppose that **E** is directed along the y axis and **H** along the z axis as in article 77. The component velocities of an ion due to the circular motion about the lines of magnetic force relative to the moving axes XYZ are

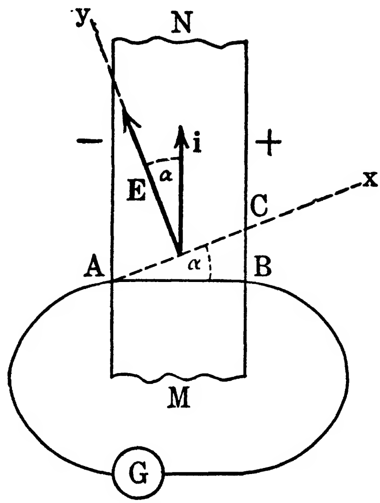


FIG. 157

$$V_x = -V \sin(\omega t + \theta) = V \sin\left(\frac{eH}{m}t - \theta\right),$$

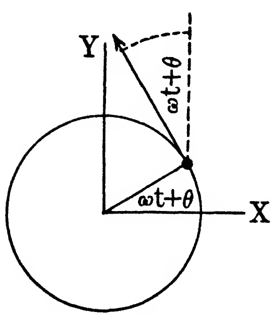


FIG. 158

$$V_y = V \cos(\omega t + \theta)$$
$$= V \cos\left(\frac{eH}{m}t - \theta\right),$$

from (77–12), where θ (Fig. 158) is the angle which **V** makes with the Y axis when $t = 0$. Therefore the components of velocity relative to the axes xyz fixed in the observer's inertial system are

$$\left.\begin{aligned} v_x &= u + V_x = u + V \cos\theta \sin\frac{eH}{m}t - V \sin\theta \cos\frac{eH}{m}t, \\ v_y &= V_y = V \cos\theta \cos\frac{eH}{m}t + V \sin\theta \sin\frac{eH}{m}t. \end{aligned}\right\} \quad (79\text{–}1)$$

If $\mathbf{v}_0$ is the initial velocity of the ion, it follows from (77–6) that

$$-V\sin\theta = v_{0x} - u, \qquad V\cos\theta = v_{0y}.$$

Substituting these in (79–1) and putting

$$\sin\frac{eH}{m}t = \frac{eH}{m}t, \qquad \cos\frac{eH}{m}t = 1 - \frac{1}{2}\left(\frac{eH}{m}\right)^2 t^2,$$

since t is small, we get

$$v_x = v_{0x} + v_{0y}\left(\frac{e}{m}\right)Ht + \frac{1}{2}\left(\frac{e}{m}\right)^2 EHt^2,$$

$$v_y = v_{0y} - v_{0x}\left(\frac{e}{m}\right)Ht + \left(\frac{e}{m}\right)Et.$$

First we find the mean velocity of an ion between successive collisions. If T is the time taken to describe a mean free path, the x and y components of the mean velocity are respectively

$$\frac{1}{T}\int_0^T v_x dt, \qquad \frac{1}{T}\int_0^T v_y dt.$$

Next we average with respect to the initial velocities. As the electrons may be assumed to start out indiscriminately in all directions after collision with the fixed atoms in the conductor, the average values of v_{0x} and v_{0y} are zero. Therefore the average components of velocity are

$$\bar{v}_x = \frac{1}{6}\left(\frac{e}{m}\right)^2 EHT^2,$$

$$\bar{v}_y = \frac{1}{2}\left(\frac{e}{m}\right)ET.$$

Consequently, if n is the number of free electrons per unit volume, the current densities in the x and y directions are

$$\left.\begin{aligned} j_x &= ne\bar{v}_x = \frac{1}{6}\frac{ne^3}{m^2}EHT^2, \\ j_y &= ne\bar{v}_y = \frac{1}{2}\frac{ne^2}{m}ET. \end{aligned}\right\} \quad (79\text{–}2)$$

The resultant current, therefore, makes an angle α with the electric field given by

$$\tan\alpha = \frac{j_x}{j_y} = \frac{1}{3}\frac{e}{m}HT = \frac{1}{3}\frac{e}{m}\left(\frac{l}{\bar{v}}\right)H, \qquad (79\text{–}3)$$

if we replace T by the quotient of the mean free path l by the mean speed $\bar{v}$. Consequently the resultant electric intensity, represented by $\mathbf{E}$ in Fig. 157, is not parallel to the axis of the strip. The transverse component of the electric field is due to charges on the edges of the strip caused by electrons which have been deflected by the magnetic field.

If E_t is the transverse component of the electric intensity,

$$E_t = E\sin\alpha \doteqdot E\tan\alpha,$$

since α is small. Hence

$$\begin{aligned} E_t &= \frac{1}{3}\frac{e}{m}ETH \\ &= \frac{2}{3ne}jH, \end{aligned} \qquad (79\text{–}4)$$

as we need not distinguish here between j_y and the total current j. This equation is usually written

$$E_t = \pi_H jH, \qquad (79\text{–}5)$$

the constant π_H being known as the *Hall coefficient*. According to the simple electron theory outlined above, this coefficient has the value

$$\pi_H = \frac{2}{3ne}, \qquad (79\text{–}6)$$

and, as the charge on the electron is negative, the Hall coefficient should always be negative. This would mean that the angle α should be laid off in the opposite sense to that illustrated in Fig. 157. Experiment shows, however, that the Hall coefficient is positive in nearly as many cases as it is negative. While a more rigorous development of the electron theory of the Hall effect leads to a slightly different numerical coefficient than

that of (79–6), the sign remains unchanged. No satisfactory explanation of the positive Hall coefficient has been proposed.

In terms of the Hall coefficient the deflection of the equipotential lines is given by

$$\tan \alpha = \frac{E_t}{E} = \pi_H \left(\frac{j}{E}\right) H = \frac{\pi_H}{\rho} H, \qquad (79\text{–}7)$$

where ρ is the resistivity.

Values of the Hall coefficient for some common metals at 18° C are given in the table. The coefficient varies considerably with temperature, and even reverses sign in certain cases. The reader will find a detailed account of this effect and other related effects in L. L. Campbell: *Galvanomagnetic and Thermomagnetic Effects*.

Metal	π_H (e.m.u.)
Ag	$-8.9(10)^{-4}$
Au	-7.2
Cu	-5.3
Al	-3.9
Zn	10.
Fe	10.
Ni	$-28.$

Problem 79a. The resistivity of gold is 2420 e.m.u. Find the deflection of the potential lines in a gold strip through which a current is passing when a normal magnetic field of 10,000 gauss is applied. Ans. $-12'.5$.

80. Space Charge. — In our previous treatment of the motion of ions in external electric and magnetic fields we have neglected the forces exerted by the ions on one another. This procedure is justifiable if ions of both signs are present in equal numbers so that attractions and repulsions balance, or, in the case of ions of one sign, if the ions are few enough and far enough apart so that the forces they exert on one another are negligibly small. In some cases, however, such as that of the hot filament vacuum tube, we have to deal with a dense current of ions of

one sign (art. 66) without the presence of compensating ions of the opposite sign. Therefore we must investigate the electric field due to the charged ions themselves in the space through which a current of ions is passing. We shall assume the presence of an external electric field only.

Let ρ be the charge per unit volume of the ions and V the electric potential in the region through which they are passing. Then Poisson's equation (26–1) must be satisfied. As $\kappa = 1$ in the empty space under consideration,

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -4\pi\rho, \tag{80–1}$$

both V and ρ being expressed in electrostatic units. Next we have from the definition of the current density j,

$$j = \rho v, \tag{80–2}$$

where v is the velocity of the ions, and finally the law of conservation of energy requires that

$$\tfrac{1}{2}mv^2 + eV = C \text{ (constant)}. \tag{80–3}$$

If we eliminate ρ and v from these equations we are left with a relation between the current density j and the potential for the case of a current consisting of ions of one sign only.

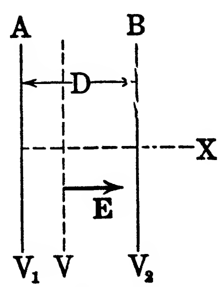


FIG. 159

We shall consider two typical cases; first a current of ions passing from one plate of a parallel plate condenser to the other, and second a current passing from a heated filament to a coaxial cylindrical electrode.

Parallel Plates. — Let the two plates A and B be at potentials V_1 and V_2 (Fig. 159) respectively. Take the X axis at right angles to the plates and suppose that the ions come off from A with negligible initial velocities. As the potential is a function of x only, (80–1) becomes

$$\frac{d^2 V}{dx^2} = -4\pi\rho. \tag{80–4}$$

Since E equals the space rate of decrease of V, this equation tells us that the electric intensity must change rapidly with x in a region in which the charge density ρ is large.

As $V = V_1$ when $v = 0$, the constant C in (80–3) is eV_1 and

$$\tfrac{1}{2}mv^2 = e(V_1 - V). \tag{80–5}$$

Eliminating ρ and v from (80–2), (80–4) and (80–5),

$$\frac{d^2V}{dx^2} = -4\pi j\sqrt{\frac{\dfrac{m}{2e}}{V_1 - V}}. \tag{80–6}$$

Multiplying by

$$2\frac{dV}{dx}dx = 2dV$$

and integrating,

$$\left(\frac{dV}{dx}\right)^2 - \left(\frac{dV}{dx}\right)_1^2 = 16\pi j\sqrt{\frac{m}{2e}}\,(V_1 - V).$$

The large number of ions in the neighborhood of the plate A neutralizes the field near the plate, so that

$$E_1 = -\left(\frac{dV}{dx}\right)_1 = 0.$$

We have then

$$\frac{dV}{dx} = -4\sqrt{\pi j}\left\{\frac{m}{2e}(V_1 - V)\right\}^{1/4},$$

where we have taken the minus sign before the radical since E is directed along the X axis and therefore the potential falls as we pass from A to B. Integrating again and solving for the current density,

$$j = \frac{1}{9\pi}\sqrt{\frac{2e}{m}}\frac{(V_1 - V)^{3/2}}{x^2}.$$

If we make x equal to the distance D between the plates, $V_1 - V$ becomes the total potential drop $V_0 \equiv V_1 - V_2$. Hence

$$j = \frac{1}{9\pi}\sqrt{\frac{2e}{m}}\frac{V_0^{3/2}}{D^2}, \tag{80–7}$$

all the quantities in this equation being expressed in electrostatic units.

Generally the ions giving rise to space charge are electrons and consequently both e and V_0 are negative, the plate B being at higher potential than A. Taking account of the fact that $V_1 - V$ in (80–6) is negative in this case, we find

$$j = -\frac{1}{9\pi}\sqrt{-\frac{2e}{m}}\,\frac{(-V_0)^{3/2}}{D^2} \quad . \tag{80–8}$$

If we substitute the value of e/m given in (77–17) the current density j_p in amperes per square centimeter is

$$j_p = -\,2.33(10)^{-6}\frac{V_p^{\,3/2}}{D^2}, \tag{80–9}$$

where V_p is the excess of the potential of B over A expressed in volts.

The right-hand side of the equation represents the maximum current density that can be obtained for a given V_0 and D. Until this maximum is reached the current increases with increasing rate of production of ions at A. After the maximum current has been obtained, however, no further increase in the rate of production of ions has any effect on the current so long as the potential drop V_0 remains unchanged, the current being limited by the *space charge* due to the ions between the two plates. The maximum current does not obey Ohm's law, being proportional to the three-halves power of the potential drop instead of the first power.

Coaxial Cylinders. — This is the case of a heated filament F (Fig. 160) emitting ions which move radially through an evacuated space to a cylindrical receiving electrode C. Let a be the radius of the filament and b that of the outer electrode. Using polar coordinates in a plane perpendicular to the common axis of the cylinders, it is clear that the potential V is a function of the radius vector r alone. Remembering that the left-hand side of Poisson's equation is the same

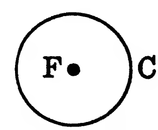


FIG. 160

as that of Laplace's equation, we have from (27–7)

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dV}{dr}\right) = -4\pi\rho, \tag{80–10}$$

provided we restore the factor $1/r^2$ which was lost by division. As the current is radial we shall use j_l to represent the current per unit length of the cylinders. Then

$$j_l = 2\pi r\rho v. \tag{80–11}$$

Finally, if V_1 is the potential of the filament we have as before

$$\tfrac{1}{2}mv^2 = e(V_1 - V). \tag{80–12}$$

Eliminating ρ and v from these three equations,

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dV}{dr}\right) = -\frac{2j_l}{r}\sqrt{\frac{\frac{m}{2e}}{V_1 - V}},$$

or, putting U for $V_1 - V$,

$$r\frac{d^2U}{dr^2} + \frac{dU}{dr} = 2j_l\sqrt{\frac{\frac{m}{2e}}{U}}.$$

On integration this gives

$$j_l = \frac{2}{9}\sqrt{\frac{2e}{m}}\frac{U^{3/2}}{r\beta^2},$$

where

$$\beta \equiv \log\frac{r}{a} - \frac{2}{5}\left(\log\frac{r}{a}\right)^2 + \frac{11}{120}\left(\log\frac{r}{a}\right)^3 - \cdots.$$

Making U equal to the potential drop V_0 between the two electrodes and r equal to the radius b of the outer cylinder, the maximum current is found to be

$$j_l = \frac{2}{9}\sqrt{\frac{2e}{m}}\frac{V_0^{3/2}}{b\beta^2}, \tag{80–13}$$

the quantities involved being measured in electrostatic units. Dividing by the circumference $2\pi b$ of the outer electrode or

plate the current per unit area is found to be

$$j = \frac{1}{9\pi}\sqrt{\frac{2e}{m}}\,\frac{V_0^{3/2}}{b^2\beta^2}.$$

If a and b are both large, the difference between them being small,

$$\beta = \log\frac{b}{a} = -\log\left(1 - \frac{b-a}{b}\right) \doteq \frac{b-a}{b},$$

to a sufficiently close approximation, and

$$j = \frac{1}{9\pi}\sqrt{\frac{2e}{m}}\,\frac{V_0^{3/2}}{(b-a)^2},$$

in agreement with (80–7).

Actually the ions are electrons here and we have

$$j_l = -\frac{2}{9}\sqrt{-\frac{2e}{m}}\,\frac{(-V_0)^{3/2}}{b\beta^2}, \qquad (80\text{–}14)$$

as in the case of (80–8). The current per unit length in amperes is given in terms of the potential difference V_p in volts by

$$j_{lp} = -14.65(10)^{-6}\frac{V_p^{3/2}}{b\beta^2}. \qquad (80\text{–}15)$$

Problem 80a. If $a = 1$ mm, $b = 10$ mm, what is the maximum current that can be obtained from the filament of a cylindrical vacuum tube under a potential of 12 volts? ($\beta = 0.97$) Ans. 0.65 milliampere/cm.

81. The Discharge Tube. — The study of the passage of electric currents through gases at low pressure has revealed many of the properties of gaseous ions and has led to the discovery of the electron. The gas is usually contained in a glass tube G (Fig. 161) provided with metal electrodes A and C by which the current enters and leaves the tube.

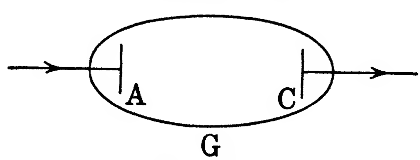


Fig. 161

In a gas at atmospheric pressure there are normally present several ions per cubic centimeter. This number may be greatly

augmented by *ionizing* the gas through the agency of X-rays or rays from a radio-active substance, which eject electrons from neutral molecules, forming pairs of oppositely charged ions. While many of the ejected electrons remain free at very low pressures, they generally become attached to neutral molecules at higher pressures, forming negative ions of molecular dimensions. The ions of opposite sign attract one another and tend to recombine into neutral molecules. If no impressed field is present a state of equilibrium is speedily reached, in which recombination goes on at the same rate as the production of new ions by the ionizing agent.

Consider a gas at a pressure of a few millimeters of mercury in a discharge tube, the gas being subject to a constant ionizing agent. If a potential drop V exists between the electrodes A and C, positive ions are driven by the field toward the cathode C and negative ions toward the anode A, and those which do not recombine on the way ultimately reach the electrodes, giving rise to a current i through the tube. As the field is increased, the ions are swept more rapidly toward the electrodes and therefore have less opportunity to recombine before reaching them. Consequently the current increases until the field becomes so great that all the ions produced reach the electrodes, when a condition of *saturation*, represented by s on the current curve of Fig. 162, is attained. Further increase of V gives rise to no additional current until the point c is reached when a very rapid growth of current takes place. This increase is due to the fact that the field has become so great that an ion acquires sufficient energy between successive collisions to eject an electron from the next neutral molecule which it strikes. This process, known as *ionization by collision*, gives rise to a greatly increased number of ions in the gas, and consequently the current grows very rapidly, the discharge assuming the nature of a *spark*.

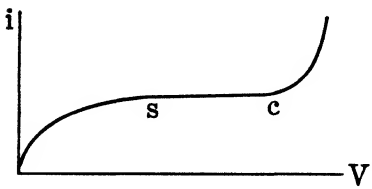


FIG. 162

When a gas is not subject to an applied ionizing agent the only current at low potentials is that due to the few ions normally present. Not until the field has been increased sufficiently to produce ionization by collision is there an appreciable current. The field necessary to produce a discharge, however, varies greatly with the pressure, as indicated in Fig. 163 for air. At pressures

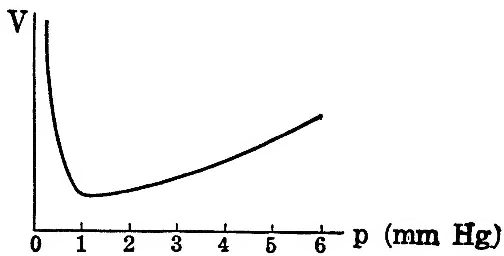


FIG. 163

of several millimeters of mercury the discharge is easily produced, but as the pressure is decreased below a millimeter or thereabout the potential required increases very rapidly. In highly evacuated tubes, therefore, it is necessary to furnish a supply of ions in order to cause a current to pass. This is generally done by using a heated filament for cathode, as in the vacuum tube of article 80, the electrons emitted by the filament carrying negative charge from cathode to anode, and therefore giving rise to a current from anode to cathode.

We shall now describe in more detail the nature of the discharge between cold electrodes. The spark observed at atmospheric pressure, which requires a potential of some hundreds or thousands of volts, becomes broadened out as the pressure is reduced, until, at pressures of a few millimeters of mercury, a steady discharge is obtained of a quite different character from that at atmospheric pressure. The luminosity, which is observed

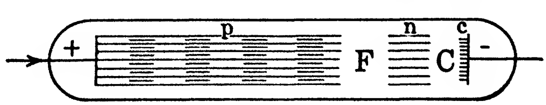


FIG. 164

on the surface of the cathode at higher pressures, stretches out into the tube, separating into the *negative glow n* (Fig. 164) and the *cathode glow c*, the two being separated by a dark region *C* known as the *Crookes dark space*. Similarly the glow on the anode extends out into the tube, becoming striated under suitable conditions. This region of luminosity, *p*, is known as the *positive column*, and is separated from the negative glow by the *Faraday dark space F*. If the length of the tube is increased, the pressure of the gas remaining constant, the positive column extends so as to fill the additional space, the other regions remaining unchanged in dimensions. As the current through the tube is increased, the cathode glow extends further over the surface of the cathode, the potential difference between the electrodes remaining nearly constant until the entire surface of the cathode is covered. Investigation of the field in the tube by means of exploring electrodes indicates that most of the drop in potential takes place near the surface of the cathode, the electric intensity varying from point to point somewhat as shown in Fig. 165.

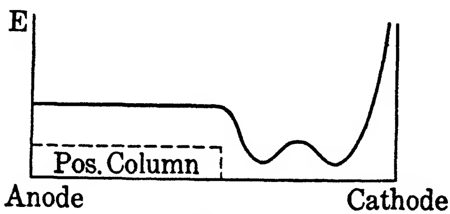


FIG. 165

The current is initiated by the collision of ions already present with neutral molecules, the negative ions produced by collision at low pressures being largely free electrons and the positive ions charged molecules. The negative ions, on account of their smaller mass, have a much greater mobility than the positive ions. Consider first the positive ions. Moving toward the cathode, they acquire greater kinetic energy between successive collisions close to the cathode than elsewhere in the tube, as the field is most intense there. Consequently most of the ionization produced by impact of the positive ions take place at the surface of the cathode. The negative ions formed there are swept away more quickly than the positive ions on account of their

greater mobility. This action results in the formation of a positive space charge in the region close to the cathode, which explains the rapid fall in potential and large electric field.

The negative ions formed near the cathode acquire sufficient kinetic energy in passing through the Crookes dark space to ionize by collision in the negative glow, thus producing the luminosity found there. The striations existing in the positive column are probably separated by distances just sufficient for the ions to acquire an energy that will enable them to excite if not to ionize the neutral molecules with which they collide. In this way the periodic luminosity in the positive column is explained.

At very low pressures many of the electrons liberated from the cathode by impact of positive ions travel the whole length of the tube without colliding with gaseous molecules. These streams of electrons are the *cathode rays* (art. 77) which J. J. Thomson deflected by perpendicular electric and magnetic fields in his determination of the ratio of charge to mass. When these electrons strike the walls of the tube, or a metallic *anti-cathode* placed in their path, their sudden stoppage gives rise to the electromagnetic waves known as *X-rays*.

82. The Magnetron. — In article 80 we discussed the ion current passing from a straight filament to a surrounding cylindrical electrode under the action of a radial electric field. We shall now investigate the characteristics of such a tube when a uniform magnetic field parallel to the filament is added. This device is known as a *magnetron*. We have here a case of crossed fields, but whereas the magnetic field is uniform the electric field is of a more complicated type. If no ions were present the electric intensity would vary inversely with the distance r from the axis as in the cylindrical condenser. Space charge distorts the electric field, however, so that no such simple relation holds other than as a first approximation. Fortunately the important characteristics of the magnetron are independent of space charge.

We shall suppose that the electrons emitted by the hot filament *AB* (Fig. 166) come off with negligibly small initial velocities. We shall neglect the magnetic field produced by the heating current in the filament, taking into account only the external magnetic field parallel to the axis of the tube and the radial electric field due to the difference in potential between the filament and the surrounding cylindrical electrode. It will be

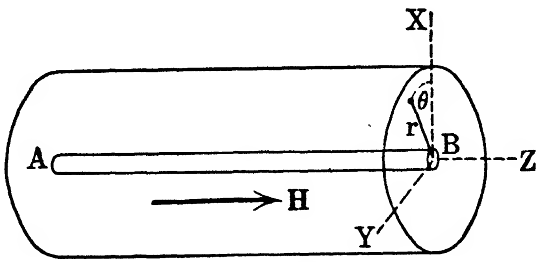


FIG. 166

convenient to use cylindrical coordinates r, θ, z, the last being measured parallel to the axis of the tube. Let a denote the radius of the filament and b that of the outer electrode. Then, as the magnetic field does no work, the energy equation is

$$\tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) = e\int_a^r E dr, \qquad (82\text{–}1)$$

the dot over a letter indicating differentiation with respect to the time.

The torque about the axis is due to the magnetic field alone and is equal to $-er\dot{r}H$. Equating this to the rate of increase of angular momentum,

$$m\frac{d}{dt}(r^2\dot{\theta}) = -er\dot{r}H,$$

and integrating,

$$r\dot{\theta} = -\frac{e}{mr}\int_a^r H r dr. \qquad (82\text{–}2)$$

If, now, we use (82–2) to eliminate $r\dot{\theta}$ from (82–1),

$$\dot{r}^2 = 2\frac{e}{m}\int_a^r E dr - \frac{e^2}{m^2r^2}\left[\int_a^r H r dr\right]^2. \qquad (82\text{–}3)$$

Since the second term on the right increases with increasing r more rapidly than the first, there will be some value of r for which $\dot{r}$ vanishes. At this distance from the filament, the electrons, deflected from their initially radial paths by the magnetic field, have reached the extreme limit of their radial motion and are starting back toward the filament. If this critical value of r is less than b, no electrons reach the outer electrode. By increasing H the critical distance is made smaller. Therefore, if we gradually augment the magnetic field parallel to the axis of a given tube, keeping the potential drop constant, the current is suddenly cut off when a critical value of H is reached. To find the condition for cut-off we make $r = b$ and put $\dot{r} = 0$ in (82–3). Now the potential drop between the electrodes of the tube is

$$V = \int_a^b E dr,$$

and the magnetic flux through the tube is

$$N = 2\pi \int_a^b H r dr.$$

Consequently the condition for cut-off is

$$V = \frac{eN^2}{8\pi^2 m b^2}. \tag{82–4}$$

If a good vacuum is maintained in the tube, the cut-off is quite sharp, as illustrated in the experimental curve of Fig. 167, showing the current i plotted against H for constant V. It is important to note that the formula (82–4) for cut-off is independent of space charge, since only the total potential drop is involved.

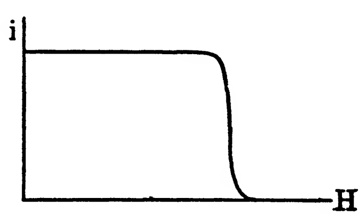


FIG. 167

The magnetron provides a means of cutting off a current by means of a magnetic field. It may also be used to measure the strength of a uniform magnetic field by placing it with its axis

parallel to the lines of force and adjusting the potential V until cut-off occurs. Finally the tube may be used to obtain the value of e/m for the electron by adjusting known electric and magnetic fields so as to secure the condition of cut-off.

Problem 82a. A magnetron consists of a filament of 0.2 mm radius surrounded by a cylindrical electrode of 3 cm radius. It is observed that when the tube is placed with its axis parallel to the lines of force of a uniform magnetic field, cut-off takes place when the potential on the tube is 8.85 volts. Find H. Ans. 6.66 gauss.

83. Diamagnetism and the Zeeman Effect. — On the electron theory the Ampèrian circuits responsible for the magnetic properties of matter are supposed to be due to rings of electrons revolving about the nuclei of the atoms. Consider one such circular ring; suppose it to consist of N electrons of charge e revolving in a circle of radius a with angular velocity ω under a force F directed toward the center of the circle. As a charge Ne passes each point of the circumference in a time $2\pi/\omega$, the current is

$$i = \frac{Ne\omega}{2\pi}, \tag{83-1}$$

provided e is expressed in electromagnetic units, and the magnetic moment of the ring (art. 70) is

$$M = iA = \tfrac{1}{2}Ne\omega a^2. \tag{83-2}$$

In general an atom or molecule contains several such rings, the resultant magnetic moment being the vector sum of the magnetic moments of the individual rings. If the resultant magnetic moment is not zero, the atom or molecule has the properties of a small magnet and the medium of which it is a constituent is paramagnetic or ferromagnetic. In a great many cases, however, the resultant magnetic moment vanishes, and then the atom or molecule is subject to no torque when placed in an external magnetic field. In all cases, however, the application of an external field changes the magnetic moments of the rings of electrons, diminishing the moments of those whose axes

lie in the direction of the field and increasing the moments of those whose axes are opposite to the field. This effect, known as *diamagnetism*, produces a resultant magnetic moment opposite to the applied field. It probably exists in all substances, although in paramagnetic and ferromagnetic media it is masked by the much greater opposite effect due to the orientation of the magnetic atoms or molecules by the field.

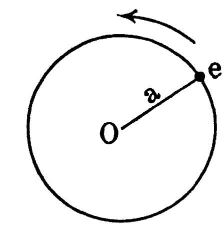


FIG. 168

To investigate the effect of an external magnetic field on the magnetic moment of a ring of electrons consider the motion of a single electron (Fig. 168). In the absence of a magnetic field the only force on the electron is the central force F, and its angular velocity ω_0 is given by the equation of motion

$$m\omega_0^2 a = F. \tag{83-3}$$

If, now, a uniform magnetic field H_1 is applied at right angles to the plane of the figure, the lines of force being directed toward the reader, a magnetic force

$$evH_1 = e\omega a H_1$$

directed away from O is brought into being. Consequently the equation of motion is now

$$m\omega^2 a = F - e\omega a H_1. \tag{83-4}$$

Eliminating F by means of (83-3),

$$\omega^2 - \omega_0^2 = -\frac{eH_1}{m}\omega.$$

If we put $\Delta\omega$ for the increase $\omega - \omega_0$ in the angular velocity,

$$\begin{aligned}\Delta\omega &= -\frac{\omega}{\omega+\omega_0}\frac{eH_1}{m}\\ &= -\frac{eH_1}{2m}\end{aligned} \tag{83-5}$$

to a sufficient degree of approximation, since ω differs very little from ω_0 even in the strongest fields available in the laboratory.

From (83–2) and (83–5) we find for the change in the magnetic moment of the ring

$$\Delta M = \tfrac{1}{2}Nea^2\Delta\omega = M\frac{\Delta\omega}{\omega}$$

$$= -\frac{MeH_1}{2m\omega}. \tag{83–6}$$

Therefore the magnetic moment of a ring of electrons whose axis lies in the direction of the field, as in the case under discussion, is decreased by the field. If the magnetic axis had been directed opposite to H_1, the effect of the field would have been to increase the moment by an equal amount. Consider an atom which contains two rings of electrons of equal moments oppositely oriented, so that in the absence of a magnetic field it has zero resultant moment. The effect of a field, parallel to the axis of the first ring, is to decrease the moment of this ring and increase that of the oppositely oriented ring. Therefore a magnetic moment opposite to the field is imparted to the atom.

If the plane of the ring is not at right angles to the field, it is evident that only the component of H_1 parallel to M is effective in altering the angular velocity and the magnetic moment of the ring. If, then, the angle between M and H_1 is denoted by θ,

$$\Delta M = -\frac{MeH_1}{2m\omega}\cos\theta, \tag{83–7}$$

and the component of the added magnetic moment in the direction of the field is

$$\Delta M\cos\theta = -\frac{MeH_1}{2m\omega}\cos^2\theta.$$

To find the intensity of magnetization in a diamagnetic medium, we must sum up over all the n rings in a unit volume. Assuming that the axes of the rings are directed at random, the number making angles between θ and $\theta + d\theta$ with H_1 is

$$dn = \frac{2\pi\sin\theta d\theta}{4\pi}n = \tfrac{1}{2}n\sin\theta d\theta,$$

and

$$I = \int \Delta M \cos \theta dn = -\frac{nMeH_1}{6m\omega}$$
$$= -\frac{nNe^2a^2}{12m} H_1. \qquad (83\text{–}8)$$

The field H_1 in (83–8) is not the mean field H in the medium but rather the field external to the atom under consideration. It corresponds to the E_1 of article 17. By analogy with (17–3) we have $H_1 = H + (4\pi/3)I$.

As nN is just the total number of electrons per unit volume, (83–8) can be used to calculate a mean radius of the atom from measured values of the susceptibility of diamagnetic substances, or, if a is taken as known, the number of electrons revolving in rings can be computed. The results are in fair accord with expectations.

The phenomenon of diamagnetism constitutes the strongest kind of evidence for Ampère's theory of magnetism. For if magnetic properties were due to the presence of actual magnetic dipoles in the atom instead of Ampèrian circuits it is difficult to see how a medium could be diamagnetic. In the electrostatic analog, for instance, where true dipoles are responsible for the properties of the medium, we never find a value of the specific inductive capacity less than unity.

As light consists of electromagnetic waves (Chapter XVI), the change in angular velocity given by (83–5) is responsible for a change in the frequency ν of light emitted by an electron revolving about the nucleus of an atom placed in an external magnetic field H. If the revolution is in a plane perpendicular to the lines of force,

$$\Delta\nu = \frac{\Delta\omega}{2\pi} = \pm \frac{eH}{4\pi m}, \qquad (83\text{–}9)$$

the upper or the lower sign being taken according as the magnetic axis of the ring is opposite or parallel to H. This influence of a magnetic field on the frequency of light is known as the *Zeeman effect.*

Now in order to account for the monochromatic character of the light emitted from a radiating gaseous atom, we must suppose that the emitting electron is bound to the atom by a simple harmonic force of restitution. Its motion, even if it is describing a circular orbit, may be treated as the resultant of linear simple harmonic vibrations. For simplicity let us confine our attention to one such linear simple harmonic motion. To analyse the effect of an impressed magnetic field we shall resolve this vibration into two linear simple harmonic vibrations, one in the direction of the field and the other in a plane perpendicular to the field. Since linear simple harmonic motion may be represented as the projection on a diameter of uniform motion in a circle, the latter may be resolved into two circular motions in opposite senses in the plane perpendicular to H. Along one diameter the displacements combine to give rise to a simple harmonic oscillation of amplitude double the common radius of the two circular motions, whereas along a diameter at right angles the displacements due to the circular motions annul each other. Altogether, then, we have replaced the original linear oscillation by three component oscillations: (1) a linear vibration along the lines of force, (2) a circular motion about the lines of force in the positive sense and (3) a circular motion about the lines of force in the negative sense. Let us denote by ν_0 the frequency in the absence of the field and by ν_1, ν_2 and ν_3 the frequencies of the three component motions when the field is applied. As a magnetic field exerts no force on a charge moving parallel to the lines of force, the frequency of the linear vibration (1) is unaffected by the field. On the other hand, that of (2) is decreased by the amount specified by (83–9) and that of (3) is increased by an equal amount. Therefore we have the three frequencies

$$\nu_2 = \nu_0 - \frac{eH}{4\pi m}, \qquad \nu_1 = \nu_0, \qquad \nu_3 = \nu_0 + \frac{eH}{4\pi m}, \quad (83\text{–}10)$$

when the field is applied, in place of the single frequency ν_0. Actually, since e is negative, ν_3 is the smallest and ν_2 the largest of these frequencies.

There are two principal ways of viewing the radiation from a source placed in a field. If the slit of the spectroscope is placed at a point on a line drawn from the source in the direction of the lines of force we have *longitudinal observation*, whereas if the light is viewed from a point on a line drawn from the source at right angles to the lines of force we have *transverse observation*. It will be shown in article 137 that the radiation from an oscillating charge is greatest in directions at right angles to the oscillation and zero in the line of the oscillation. Furthermore the electric vector, the direction of which determines the polarization of the emitted light, lies in the plane determined by the oscillation. Therefore in longitudinal observation ν_2 and ν_3 appear circularly polarized and no light of frequency ν_1 is present. On the other hand, all three frequencies appear as plane polarized light in transverse observation, ν_2 and ν_3 with the electric vector perpendicular to H (s polarization) and ν_1 with the electric vector parallel to H (p polarization). These three frequencies constitute the *normal Lorentz triplet*. One of the earliest methods of calculating the ratio e/m of the electron was from measurements of the separation found in the Zeeman effect.

The analysis of the Zeeman effect given here fits the observations in every respect for simple spectra. In the majority of cases, however, the resolution is more complex than that predicted by this simple theory, and for an adequate explanation the concept of the spinning electron and the quantum theory are necessary.

Problem 83a. On Ampère's theory of magnetism the mean field inside a magnetic medium is B, not H. Show, nevertheless, that the mean magnetic intensity to which an atom is subject is $H + (4\pi/3)I$ just as in the case of the electrostatic analog. (Consult art. 17 and art. 75.)

Problem 83b. From (83–8) find the susceptibility and the permeability of a diamagnetic medium in terms of $\sigma \equiv \dfrac{nNe^2a^2}{12m}\cdot$

$$\text{Ans. } \epsilon = -\frac{\sigma}{1+\frac{4\pi}{3}\sigma}, \qquad \mu = \frac{1-\frac{8\pi}{3}\sigma}{1+\frac{4\pi}{3}\sigma}.$$

84. Unipolar Induction. — Consider a symmetrically magnetized steel cylinder SN (Fig. 169) arranged so that it can rotate about its axis and provided with a side arm AD the end D of

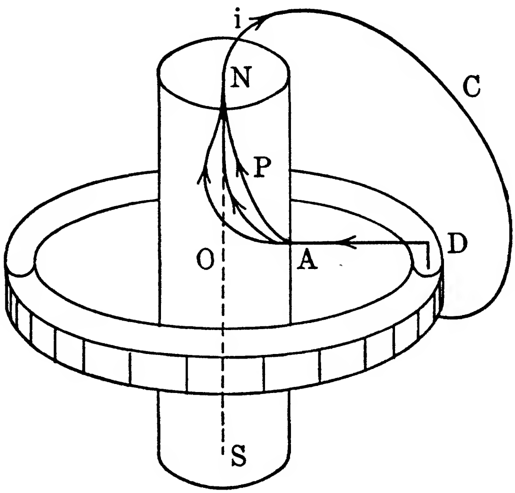


FIG. 169

which dips into a coaxial trough of mercury. The trough and the portion NCD of the circuit are held fixed. If a current passes from the trough through the side arm into the magnet, emerging along the axis at N, the magnet experiences a torque which causes it to rotate in the positive sense about the axis SN. We have, here, a simple type of motor.

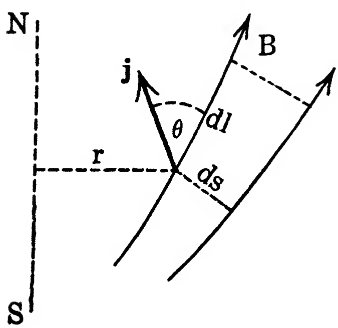


FIG. 170

To calculate the torque on the current passing through the side arm and magnet, we have for the force per unit volume

$$\mathbf{F}_r = \mathbf{j} \times \mathbf{B},$$

from (74–5).

If θ (Fig. 170) is the angle between $\mathbf{j}$ and $\mathbf{B}$ and ϕ the azimuth measured around the axis SN of the magnet, the force on a

volume element $rd\phi dsdl$ is

$$jB \sin \theta rd\phi dsdl$$

at right angles to the plane of the figure. We are measuring dl along and ds at right angles to the lines of induction. The torque due to this force is

$$d^2L = jB \sin \theta r^2 d\phi dsdl$$

in the positive sense about the axis SN.

Now the current di passing through the face $rd\phi dl$ of the volume element is

$$di = j \sin \theta rd\phi dl,$$

and hence

$$d^2L = rBdsdi.$$

But if dN is the flux of induction through an annular ring of width ds described about the axis of the magnet,

$$dN = 2\pi rBds,$$

since the magnet is symmetrically magnetized and ds is perpendicular to the lines of induction. Consequently

$$d^2L = \frac{didN}{2\pi}.$$

Summing up dl and $d\phi$ is equivalent to summing up di without change in N, giving

$$dL = \frac{idN}{2\pi}.$$

Finally, summing up dN gives for the total torque

$$L = \frac{iN}{2\pi}, \qquad (84\text{–}1)$$

where N is the flux of induction through the circle at right angles to the axis of the magnet of radius OD (Fig. 169). Evidently we have made no error in treating $\mathbf{j}$ as if it lay in the plane through NS since the component of $\mathbf{j}$ at right angles to this plane contributes nothing to the torque. As the lines of induction

outside the magnet have the opposite sense to those in its interior, the torque is greatest when the side arm is shortest.

The torque expressed by (84–1) is that produced by the magnetic field on the portion of the current circuit lying in the movable part of the system. Before we can conclude that this is the total torque we must show that there is no torque exerted by the current circuit on the poles of the magnet.

The spreading of the current in the interior of the magnet may be taken into account by replacing the current by a number of closed current filaments which follow the same path outside the magnet but separate in its interior. Let the circuit *NCDAPN* (Fig. 169) be one such elementary circuit, the current flowing in it being designated by di. If H_ϕ is the component of the magnetic intensity in the direction of increasing ϕ in the interior of the magnet due to this current filament and ρ the volume density of magnetic charge, the torque on an annular ring of the magnet of radius r and cross-section $ds\, dl$ such as is shown in section in Fig. 170 is

$$dL = r\rho dl ds \int_0^{2\pi} H_\phi r d\phi.$$

Now, according to Ampère's law (72–3),

$$\int_0^{2\pi} H_\phi r d\phi = 4\pi di$$

if the ring encircles the current di, or zero if it does not. Moreover, the magnetic charge in the entire ring is

$$dm = 2\pi r\rho dl ds.$$

Hence

$$dL = 2di dm.$$

To get the total torque due to the current filament we must sum this expression over all annular rings encircling the filament. But the total magnetic charge on any isolated body vanishes. Hence $\int dm = 0$ and the torque under consideration vanishes.

The magnetic charges under consideration are, of course,

those of the equivalent magnetic shells by which we are supposing each Ampèrian circuit in the steel cylinder to be replaced for the purpose of calculating the torque exerted by the current filament on the magnet. It might be supposed at first sight that our method of proof would fail if the current filament should thread an Ampèrian circuit. In such a case, however, we can divide the equivalent shell into two parts, one of which lies in an annular ring encircling the filament and the other of which lies in a ring which does not. As the total magnetic charge in each part of the shell is zero, the reasoning outlined above is valid.

We conclude, then, that (84–1) represents the total torque on the movable part of the system. This formula is well verified by experiment, and substitution of a brass cylinder in the magnetic field of a coaxial solenoid for the magnet verifies the theoretical deduction that the poles of the magnet are in no way responsible for the torque. The fact that the formula deduced on the basis of Ampère's theory of magnetism accords with the experimental measure of the torque shows the correctness of Ampère's conception of the nature of magnetism. For if the molecules of a magnetic medium owed their magnetic properties to the presence of actual dipoles instead of Ampèrian circuits, the flux N in (84–1) would be that of H instead of B, and as H may be less than a hundredth of B in iron, the torque would be only a small fraction of that actually observed.

The mechanism in Fig. 169 has been considered as a simple form of electric motor. It may also be operated as a generator. To show this, let $\mathcal{E}_0$ be the external electromotive force producing the current i and R the resistance of the circuit, and suppose that the magnet rotates with constant angular velocity ω against a frictional torque equal to (84–1). Then the rate at which work is done by the external electromotive force is

$$\mathcal{E}_0 i = Ri^2 + \frac{iN}{2\pi}\omega,$$

or

$$\left\{ \mathcal{E}_0 - \frac{N\omega}{2\pi} \right\} i = Ri^2.$$

As Ri^2 must equal the product of the total electromotive force by the current, we see that the rotation of the magnet gives rise to a back electromotive force

$$\mathcal{E} = \frac{N\omega}{2\pi}. \qquad (84\text{–}2)$$

If, now, we remove $\mathcal{E}_0$ and rotate the magnet mechanically, the electromotive force (84–2) will be generated.

Problem 84a. A current of 0.05 amp passes through a magnet as indicated in Fig. 169. The flux through the circle of radius OD is 1045 max. Compute the torque. Ans. 0.832 dyne cm.

Problem 84b. Deduce the torque (84–1) by computing the torque about SN on the portion NCD of the circuit of Fig. 169 and applying the law of action and reaction.

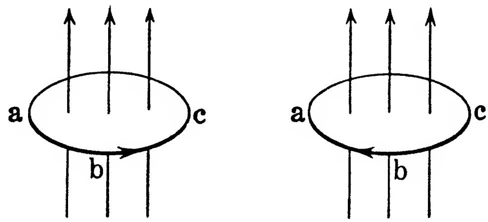
CHAPTER IX

ELECTROMAGNETIC INDUCTION

85. Faraday's Law. — Inasmuch as a current gives rise to a magnetic field, it occurred to Faraday in 1831 that a magnet might induce a current in a neighboring fixed circuit. Although he found that the proximity of a stationary magnet gives rise to no current, he noticed that a galvanometer in the circuit suffers a momentary deflection while the magnet is approaching or receding. Further investigation showed that the same effect is produced by moving a second circuit in which a current is flowing toward or away from the first, or by holding the second circuit fixed and varying the current in it. The phenomenon of current induction had been observed by Henry in America even earlier, but the discovery is generally attributed to Faraday since he was the first to publish his results.

Faraday's experiments indicate that a change in the flux of magnetic induction through a fixed circuit gives rise to an electromotive force which lasts as long as the flux is changing, the magnitude of the induced e.m.f. being proportional to the time rate of change of flux. If the flux through the circuit is decreasing, the sense of the induced electromotive force is found to be that of rotation of a right-handed screw advancing in the direction of the flux passing through the circuit, as illustrated in Fig. 171*a*, whereas if the flux is increasing, the induced electromotive force is in the opposite sense as in Fig. 171*b*. Now suppose we take the upper face of the circuit shown in the figure as positive so as to make the flux through it positive. Then, according to the convention of article 8 relating the positive sense of describing the periphery of a surface to the positive direction of the normal, the periphery is traversed in the positive sense when we pass from *a* to *b* to *c*. Consequently the induced electromotive force is positive in the case (*a*) of decreasing flux

and negative in the case (*b*) of increasing flux. As a current in the circuit would have to flow in the sense *a* to *b* to *c* to produce a flux in the same direction as that of the external field, the induced e.m.f. acts in such a sense as to produce additional flux in the *same* direction as the flux due to external sources if the



(a) Flux decreasing (b) Flux increasing

FIG. 171

latter is decreasing, and in the sense to produce additional flux in the *opposite* direction to the flux due to external sources if the latter is increasing. In each case the induced e.m.f. acts in the sense to *oppose* the change in flux which produces it.

If B is the magnetic induction, ds an element of area of any surface s having the circuit as its periphery, and γ the angle between the lines of induction and the positive normal to the surface element, the flux of induction through the surface is

$$N = \int_s B \cos \gamma ds = \int_s \mathbf{B} \cdot d\mathbf{s}.$$

Taking account of the relation between the positive face of the circuit and the positive sense of describing the periphery, *Faraday's law states that the electromotive force $\mathcal{E}$ induced by changing the flux of induction through a fixed circuit is equal to the time rate of decrease of flux, that is,*

$$\mathcal{E} = -\frac{dN}{dt}, \tag{85-1}$$

or

$$\mathcal{E} = -\int_s \frac{\partial}{\partial t}(B \cos \gamma) ds, \tag{85-2}$$

where we have to differentiate $\cos \gamma$ as well as B since the direction as well as the magnitude of the induction may change with the time.

Now let $\mathbf{B}_1$ (Fig. 172) be the magnetic induction at the time t and $\mathbf{B}_2$ that at the time $t + \Delta t$, the angle which $\mathbf{B}_1$ makes with the normal n to the surface element ds being denoted by γ_1 and that which $\mathbf{B}_2$ makes with n by γ_2. The increase in the vector $\mathbf{B}$ during the time Δt is

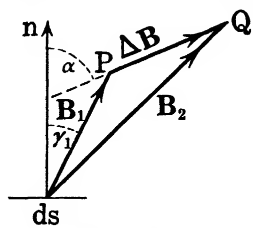


FIG. 172

$$\Delta \mathbf{B} = \mathbf{B}_2 - \mathbf{B}_1$$

in the direction PQ, and the time rate of increase of $\mathbf{B}$ at the time t is the vector

$$\frac{\partial \mathbf{B}}{\partial t} = \lim_{\Delta t = 0} \frac{\Delta \mathbf{B}}{\Delta t}$$

in the limiting direction assumed by the line PQ as Δt approaches zero.

If we represent the magnitude of the vector $\Delta \mathbf{B}$ by $|\Delta \mathbf{B}|$,

$$|\Delta \mathbf{B}| \cos \alpha = B_2 \cos \gamma_2 - B_1 \cos \gamma_1$$

from the figure, or

$$\left| \frac{\Delta \mathbf{B}}{\Delta t} \right| \cos \alpha = \frac{B_2 \cos \gamma_2 - B_1 \cos \gamma_1}{\Delta t},$$

and letting Δt approach zero

$$\left| \frac{\partial \mathbf{B}}{\partial t} \right| \cos \alpha = \frac{\partial}{\partial t} (B \cos \gamma), \tag{85-3}$$

the angle α between PQ and the normal now having its limiting value. Therefore Faraday's law becomes

$$\mathcal{E} = - \int_s \left| \frac{\partial \mathbf{B}}{\partial t} \right| \cos \alpha ds = - \int_s \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{s}. \tag{85-4}$$

In differentiating $\mathbf{B}$ with respect to the time we have used the notation for partial rather than for total differentiation for the

reason that **B** is generally a function of the coordinates as well as of the time. In (85–4) we are dealing with a fixed circuit and are concerned solely with the time rate of change of **B** at each point.

The reader must remember that $\left|\frac{\partial \mathbf{B}}{\partial t}\right|$, the magnitude of the time rate of increase of the vector **B**, is not at all the same thing as $\frac{\partial B}{\partial t}$, the time rate of increase of the magnitude of **B**. For instance, in the case of a rotating vector of constant magnitude, the latter derivative vanishes while the former does not.

As the electromotive force in this case is the line integral (48–5) of the electric intensity **E** around the circuit, we can write (85–4) in the form

$$\oint \mathbf{E}\cdot d\mathbf{l} = -\int_s \frac{\partial \mathbf{B}}{\partial t}\cdot d\mathbf{s}, \qquad (85\text{–}5)$$

provided **E** as well as **B** is expressed in electromagnetic units.

It is important to note that Faraday's law applies to any closed curve whether or not a conducting wire coincides with it, although only when a conducting circuit is present does the electromotive force give rise to a current. *In all cases Faraday's law tells us that the line integral of the electric intensity around any fixed closed curve is equal to the time rate of decrease of flux of magnetic induction through the curve.* Equation (85–5) should be compared with the circuital form (72–3) of Ampère's law, to which it is quite analogous, particularly if the latter is written in the form

$$\oint \mathbf{H}\cdot d\mathbf{l} = 4\pi\int_s \mathbf{j}\cdot d\mathbf{s}, \qquad (85\text{–}6)$$

obtained by expressing i in terms of the current density **j** over any surface s whose periphery coincides with the path along which the line integral of **H** is taken.

Faraday's law implies that the lines of electric force due to changing magnetic induction form closed curves much as do the

lines of magnetic force produced by a current in a wire. Consider, for instance, a magnet (Fig. 173) held above the plane of the paper with its axis on the perpendicular through O, the north pole being the nearer to O. If, now, the magnet is moved toward the paper, the lines of electric force in the plane of the figure are, on account of symmetry, circles with O as center, the sense of the electric intensity being that of the arrow-heads. Reversal of the direction of motion of the magnet merely reverses the sense of **E**. If a wire coincides with the line ABC, the electric field exerts a force on the free electrons in a direction everywhere tangent to the wire, and consequently a current is produced. Evidently no scalar potential exists in an electric field of this type, for such a potential would have to decrease as we pass in the direction of the electric intensity from A to B to C to A, so that after passing completely around the circuit we would come back to the starting point with a different value of the potential from that with which we had begun.

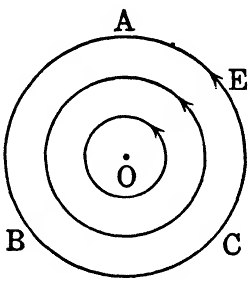


FIG. 173

In the case under consideration no part of the wire ABC is charged electrically. If, however, we cut the wire at A and insert a parallel plate condenser as in Fig. 174, the situation is altered. Now the electric field **E** due to the changing induction causes positive charge to accumulate at a and negative charge at b until the difference of potential $V_a - V_b$ between the two plates of the condenser due to these charges is equal to the electromotive force $\mathcal{E}_1$ represented by the line integral of the

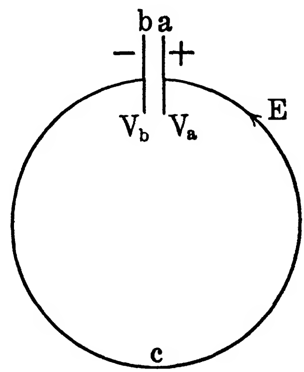


FIG. 174

electric intensity $\mathbf{E}$ along the wire from b to c to a, that is,

$$V_a - V_b = \int_{bca} \mathbf{E} \cdot d\mathbf{l} = \mathcal{E}_1.$$

For only when this condition is satisfied is the total work performed in carrying an electron from b to a through the wire equal to zero and only then can the flow of electrons stop. If the distance d between the plates of the condenser is small compared to the distance l around the circuit, $\mathcal{E}_1 = (1 - d/l)\mathcal{E}$. Hence

$$V_a - V_b = \left(1 - \frac{d}{l}\right)\mathcal{E}$$

for equilibrium.

The electric intensity in the region between the plates is not $(V_a - V_b)/d$, but rather $(V_a - V_b)/d + E$. Nevertheless the charge on the condenser, at least in the case where the specific inductive capacity of the dielectric between the plates is unity, is determined by $(V_a - V_b)/d$ alone, for the lines of force of the field E do not terminate on the plates and therefore do not lead to any additional charge.

In this connection it must be remembered that a field due to static charges can never give rise to an electromotive force around a closed circuit. For every electrostatic field possesses a potential, and therefore the line integral of the electric intensity around any closed curve vanishes, as shown in article 48. Consequently the electromotive force around a closed curve due to changing magnetic flux is unaffected by the presence of charges such as those on the condenser plates of Fig. 174.

If a circuit is located in empty space, where B and H are the same, the flux of magnetic induction through the circuit is equal to the flux of magnetic intensity and Faraday's law may be stated equally well as asserting that the induced electromotive force is equal to the time rate of decrease of the flux of magnetic intensity. In case the circuit embraces a magnetic core, as in Fig. 175, we may determine the flux either by integrating over a

surface s_1 lying entirely in empty space or by integrating over a surface s_2 cutting the core. But B lines are continuous as we pass through the surface of the core, whereas H lines are not. Therefore the flux of $H = B$ through s_1 is equal to the flux of B through s_2. Hence we see that, when the surface of integration passes through a permeable medium, the induced electromotive force must depend upon the rate of change of flux of B as expressed by (85–2) rather than upon the rate of change of flux of H, for, if the latter were the case, we would obtain different results according as we integrated over s_1 or s_2. The same conclusion may be reached from Ampère's theory of magnetism (art. 75), which makes the mean magnetic intensity inside a permeable medium equal to B.

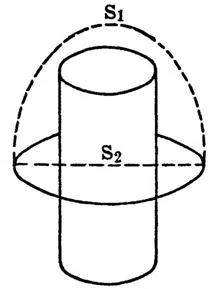


FIG. 175

Problem 85a. A long solenoid of n_1 turns per unit length and cross-section A carries an alternating current $i = i_0 \sin \omega t$. The solenoid is wound on a core of permeability μ. Find the flux of B through a closed curve which encircles the solenoid and calculate the e.m.f. around the curve.

$$\text{Ans. } N = 4\pi n_1 i_0 \mu A \sin \omega t, \qquad \mathcal{E} = -4\pi n_1 i_0 \mu A \omega \cos \omega t.$$

Problem 85b. A current $i = i_0 \sin \omega t$ flows in a very long straight wire. Calculate the flux through and e.m.f. around a rectangle of dimensions h and d which lies in a plane through the wire, the sides of length h being parallel to the wire and at distances R and $R + d$ from it.

$$\text{Ans. } N = 2 i_0 h \log_e \left(1 + \frac{d}{R}\right) \sin \omega t,$$

$$\mathcal{E} = -2 i_0 h \log_e \left(1 + \frac{d}{R}\right) \omega \cos \omega t.$$

Problem 85c. If the magnet of Fig. 173 has a moment M, is at a distance h from O, and is approaching O with a velocity v, find the flux N through a circle of radius p about O and the electric intensity E in the plane of the figure at a distance p from O.

$$\text{Ans. } N = \frac{2\pi M p^2}{(h^2 + p^2)^{3/2}}, \qquad E = -\frac{3Mphv}{(h^2 + p^2)^{5/2}}.$$

Problem 85d. The plates of the condenser of Fig. 174 are separated by a dielectric of specific inductive capacity κ. Show that the charge per unit area (neglecting edge effect) is

$$\sigma = \frac{1}{4\pi d}\{\mathcal{E}_1 + (\kappa - 1)\mathcal{E}\}.$$

Problem 85e. If

$$H_z = \frac{A}{r}\cos \sigma r \cos \omega t,$$

where $r^2 = x^2 + y^2$, what shape and size should a circuit in the XY plane have in order that the induced e.m.f. should be as great as possible? What is the maximum e.m.f.?

Ans. Circle of radius $\left(n + \frac{1}{2}\right)\frac{\pi}{\sigma}$ where n is a positive integer,

$$\mathcal{E} = \frac{2\pi\omega A}{\sigma}\sin \omega t.$$

86. Motional Electromotive Force. — In the previous article we have confined our attention to the electromotive force in a circuit which is fixed relative to the observer. Now we shall determine the electromotive force alone a conducting wire of length l which is moving with velocity v relative to the observer across the lines of force of a magnetic field. In this case we do not need to appeal to experiment, for the electromotive force is deducible at once from the expression (76–2) for the force on one of the free electrons in the moving wire. From this equation the force per unit charge due to a magnetic field is seen to be

$$\mathbf{E}_v = \mathbf{v} \times \mathbf{B},$$

where $\mathbf{v}$ represents the velocity of the wire. So, if $d\mathbf{l}$ is a vector element of length of the wire, the *motional electromotive force* is

$$\mathcal{E} = \int_l \mathbf{E}_v \cdot d\mathbf{l} = \int_l (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l} \tag{86–1}$$

in the sense of the component of $\mathbf{v} \times \mathbf{B}$ along the wire.

Consider the parallelopiped (Fig. 176) formed by the vectors $\mathbf{v}$, $\mathbf{B}$ and $d\mathbf{l}$. The vector $\mathbf{v} \times \mathbf{B}$ is equal in magnitude to the area of the base and has the direction of the altitude. Consequently

the scalar $(\mathbf{v} \times \mathbf{B})\cdot d\mathbf{l}$ is equal to the area of the base multiplied by the component of $d\mathbf{l}$ normal to the base, that is, equal to the volume of the parallelopiped. If, now, we chose the face of the

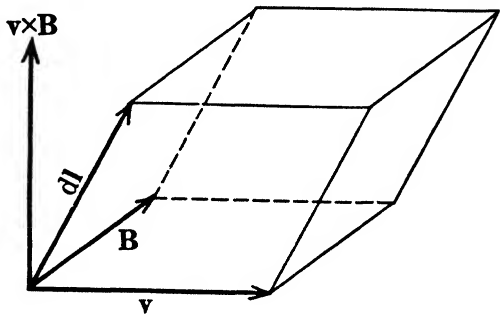


FIG. 176

parallelopiped of which $\mathbf{v}$ and $d\mathbf{l}$ are edges as base, the volume of the parallelopiped is given by $(d\mathbf{l} \times \mathbf{v})\cdot\mathbf{B} = \mathbf{B}\cdot(d\mathbf{l} \times \mathbf{v})$. Therefore

$$(\mathbf{v} \times \mathbf{B})\cdot d\mathbf{l} = \mathbf{B}\cdot(d\mathbf{l} \times \mathbf{v}),$$

and the electromotive force under consideration may be written

$$\mathcal{E} = \int_l \mathbf{B}\cdot(d\mathbf{l} \times \mathbf{v}). \qquad (86\text{–}2)$$

Now $d\mathbf{l} \times \mathbf{v}$ is the area swept over per unit time by the length $d\mathbf{l}$ of wire and $\mathbf{B}\cdot(d\mathbf{l} \times \mathbf{v})$ is the flux of induction through this area. *Therefore the motional electromotive force is equal to the flux of induction cut by the wire per unit time, or, more simply, to the number of tubes of induction cut per unit time. It acts in the direction of the component of* $\mathbf{v} \times \mathbf{B}$ *along the wire.*

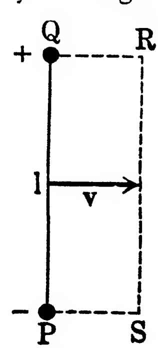


FIG. 177

Consider, for instance, a straight wire PQ (Fig. 177) of length l with a metal knob on either end moving with velocity $\mathbf{v}$ to the right through a uniform field the lines of force of which are directed perpendicular to the plane of the figure away from the reader. The electromotive force acts in the direction of the

vector $\mathbf{v} \times \mathbf{B}$, that is, from P to Q. The flux of induction cut by the wire per unit time is the flux Bvl passing through the area $PQRS$. Therefore

$$\mathcal{E} = Bvl$$

acting from P toward Q. This electromotive force urges positive electricity toward Q and negative toward P until the knobs acquire sufficient charges to produce a difference of potential equal to the electromotive force. Observe that the motional e.m.f. vanishes if either $d\mathbf{l}$ or $\mathbf{v}$ is parallel to the field, for in either of these cases no lines of force are cut.

Let us apply (86–2) to the case of a complete circuit which is moving relative to the observer. The motion may be due to the fact that the circuit has a movable arm which slides back or forth relative to the remainder of the circuit, or the circuit, while rigid, may translate or rotate relative to the observer. In any case the periphery or a part of it moves relative to the observer, having the location of the curve l_1 (Fig. 178) at the time t and that of the curve l_2 at the time $t + dt$. As we are describing the circuit in the counter-clockwise sense we must take the front face as positive. If, then, we denote by $d\mathbf{s}_l$ the vector area which is added to the circuit per unit length in the time dt, the area $PQRS$ added to the length dl in the time dt is

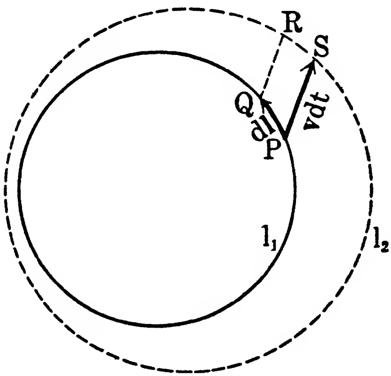


FIG. 178

$$d\mathbf{s}_l dl = -\, d\mathbf{l} \times \mathbf{v}dt,$$

the negative sign being due to the fact that the positive sense of $d\mathbf{s}_l$ is toward the reader, and (86–2) becomes

$$\mathcal{E} = -\oint \mathbf{B} \cdot \frac{\partial \mathbf{s}_l}{\partial t} dl = -\oint B \cos \gamma \left| \frac{\partial \mathbf{s}_l}{\partial t} \right| dl, \qquad (86\text{–}3)$$

the angle γ being that between **B** and the positive normal to the area $PQRS$ and the integral being taken all the way around the circuit. The integral represents the negative of the flux of B through the area added to the circuit per unit time on account of the motion of the periphery. *Therefore the motional electromotive force around a closed circuit is equal to the rate of decrease of the flux of induction through the circuit due to the motion of the periphery relative to the observer.*

Finally let us compute the total electromotive force around a closed circuit when **B** at every point is changing with the time and also the circuit is moving relative to the observer. Referring again to Fig. 178, if the circuit remains fixed in the position l_1 we have an e.m.f. equal to the rate of decrease of flux of induction through the fixed curve l_1 as specified by Faraday's law. In addition, however, we have in the case of a moving circuit a motional e.m.f. equal to the rate of decrease of flux due to the motion of the periphery. Adding the two together, *the total electromotive force $\mathcal{E}$ is equal to the total rate of decrease of flux of induction through the circuit, due in part to the change of* **B** *with the time and in part to the motion of the circuit relative to the observer.* In analytical form,

$$\mathcal{E} = -\frac{dN}{dt}, \tag{86-4}$$

where

$$N = \int_s B \cos \gamma ds = \int_s \mathbf{B} \cdot d\mathbf{s},$$

integrated over the entire area bounded by the circuit at the instant considered. This law was formulated by F. E. Neumann in 1845. It is equivalent to Faraday's law for the case of a fixed circuit, but is more general in that it applies also to a moving circuit.

Two cautionary remarks are called for by the preceding discussion. In the first place, although an e.m.f. acts around a fixed closed curve through which the flux of induction is changing whether a conducting circuit coincides with the curve or not, we cannot suppose a motional e.m.f. to exist in the absence of

moving charges, such as the electrons carried along by a moving conductor. Secondly, the statement that the motional e.m.f. along a moving wire is equal to the number of tubes of induction cut per unit time given on page 327 is correct only if the tubes of induction are considered to be stationary relative to the observer, for only in this case is the number of tubes cut equal to the flux of induction through the area swept over by the wire. The concept of moving tubes of induction is one which should be avoided as it often leads to erroneous conclusions.

A simple example of motional electromotive force is afforded by the circuit of Fig. 179. The arms *AC*, *CD*, *DE* are fixed, whereas the wire *FG* slides to the right with velocity v through a magnetic field directed perpendicular to the plane of the figure away from the reader. The positive sense of describing the circuit is *FDCG* and the e.m.f. is

$$\mathcal{E} = -\frac{dN}{dt} = -Blv, \tag{86-5}$$

where l is the width of the circuit. The negative sign indicates that the e.m.f. acts in the sense *GCDF*. It is generated entirely

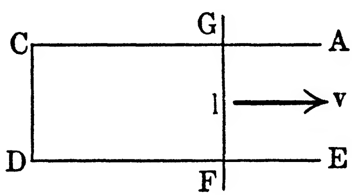


FIG. 179

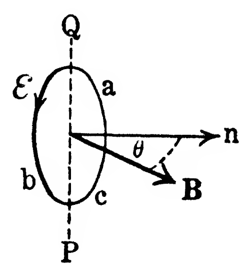


FIG. 180

in the moving wire *FG*, and the present method of computing it leads to the same result as that obtained in the case of the isolated wire of Fig. 177.

Next consider the plane coil *abc* (Fig. 180) of one turn rotating about a line *PQ* lying in its plane. We shall suppose that the circuit is in a uniform field *B* the lines of force of which are at

right angles to the axis PQ. If θ is the angle which the normal to the plane of the circuit makes with B, the flux of induction is

$$N = AB \cos \theta,$$

where A represents the area of the circuit, and the induced e.m.f. is

$$\mathcal{E} = -\frac{dN}{dt} = AB \sin \theta \frac{d\theta}{dt}. \qquad (86\text{–}6)$$

If the coil consists of n turns, each of which embraces the same flux, the e.m.f. is n times as great. Both the flux and the

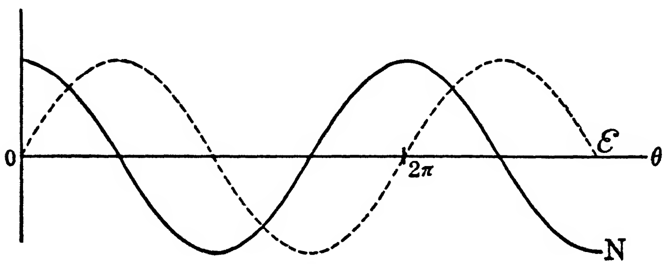


FIG. 181

induced e.m.f. reverse their senses relative to the circuit twice each revolution. If the rotation is uniform, the one is maximum when the other is zero and *vice versa*, as illustrated in Fig. 181. In this case $\theta = \omega t$, where ω is the constant angular velocity of the coil, and

$$\mathcal{E} = nAB\omega \sin \omega t. \qquad (86\text{–}7)$$

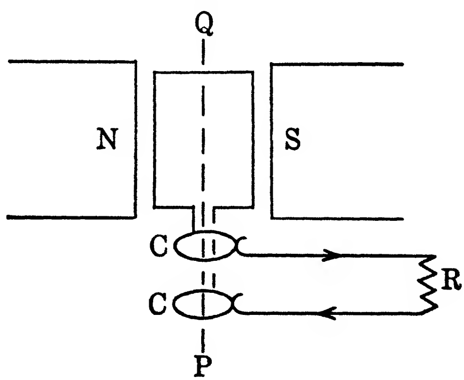


FIG. 182

The ends of the coil may be connected to *collector rings* C, C (Fig. 182), located on the axis of rotation PQ, from which the current is carried by means of brushes through an external circuit R. We have here the essen-

tials of an *alternating current generator.* The magnetic field in which the *armature coil* rotates is supplied by the *field magnet* whose poles are indicated by N and S.

If the lines of force are not perpendicular to the axis of rotation of the coil, we must understand by B in (86–6) the component of the magnetic induction at right angles to the axis, for the component of the magnetic induction parallel to the axis does not contribute to the flux through the circuit at any time during the course of a revolution.

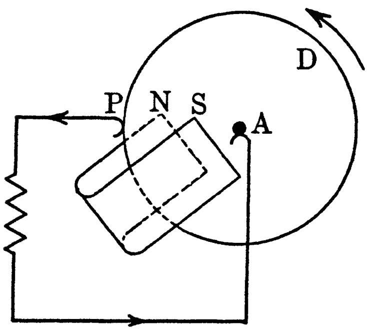


FIG. 183

A simple type of *direct current generator* making use of motional electromotive force is the *Faraday disk machine.* It consists of a copper disk D (Fig. 183) of radius a rotated with angular velocity ω between the poles N, S of a magnet, the external circuit being brought into electrical contact with the disk by brushes at the axis A and at the rim P. If, for simplicity, we assume the magnetic field to be uniform between A and P and perpendicular to the plane of the disk, (86–2) gives for the induced e.m.f.

$$\mathcal{E} = \int_0^a Bvdl = B\omega\int_0^a ldl = \tfrac{1}{2}B\omega a^2 \qquad (86\text{–}8)$$

directed from A to P. Machines of considerable power have been constructed on this principle, being known as *homo-polar generators.*

A more commonly used direct current generator is similar in construction to the alternator illustrated in Fig. 182, the collector rings being replaced by a *commutator* which reverses the sense in which the external circuit is connected to the rotating coil every half-revolution. With a single coil, as shown in the figure, this arrangement gives a unidirectional but fluctuating current. By using a number of armature coils uniformly oriented

about the axis PQ and connecting each to the external circuit only during that part of its revolution when the e.m.f. is near its maximum, an effectively constant current may be produced.

Problem 86a. Twelve Faraday disks of 10 cm radius connected in series rotate at 50 r.p.s. in fields of 10,000 gauss. Calculate the induced e.m.f. in volts. Ans. 18.8 volt.

87. Circuit with Self-Inductance and Resistance. — A current gives rise to a magnetic field and therefore contributes to the flux of induction through the circuit in which it is flowing. If the current changes, the flux is altered and an electromotive force is induced. This phenomenon is known as *self-induction.* As the flux produced by a current is positive relative to the sense in which it is flowing, the induced electromotive force is in the direction of the current if the latter is decreasing and in the opposite direction if the current is increasing. Therefore the e.m.f. of self-induction opposes any change in the magnitude of the current.

According to Ampère's law (69–6) the magnetic field and consequently the flux of induction in a paramagnetic or diamagnetic medium is proportional to the current. Hence we may write for the flux of induction through a circuit in which a current i is flowing

$$N = Li. \tag{87–1}$$

The positive coefficient L is known as the *self-inductance* of the circuit. It represents the flux of induction through the circuit due to a unit current. Provided the permeability of the surrounding medium is constant, L is a constant characteristic of the circuit.

If the current i changes, an electromotive force

$$\mathcal{E}_I = -\frac{dN}{dt} = -L\frac{di}{dt} \tag{87–2}$$

is induced in a circuit of constant self-inductance. From this equation we may define the self-inductance L as the e.m.f. induced per unit time rate of decrease of current. The practical

unit of inductance is called the *henry*. It is the self-inductance of a circuit in which an e.m.f. of one volt is induced when the current decreases at the rate of one ampere per second. The henry is equal to $(10)^9$ e.m.u. of inductance, the electromagnetic unit being the *centimeter*.

While the self-inductance of a single turn of wire is small, that of a coil of a number of closely wound turns may be quite large, especially if the wire is wound on a core of high permeability. For multiplying the turns by n multiplies the flux through each turn by n and the flux through the entire coil by n^2.

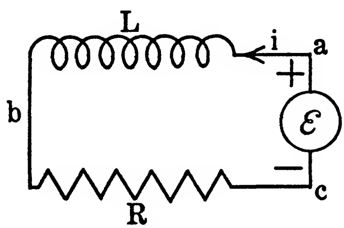


FIG. 184

Consider a circuit (Fig. 184) consisting of an impressed electromotive force $\mathcal{E}$, a coil of self-inductance L, and a resistance R connected in series. The impressed e.m.f., acting in the sense c to a, builds up a positive charge at a and a negative charge at c. If the resistance between c and a is zero the potential difference due to such charges must at every instant be equal to $\mathcal{E}$, for the resultant electric intensity must vanish in the interior of a perfect conductor. So if we denote by $\mathcal{E}_1$ the drop in potential between c and a due to the charges on the circuit,

$$\mathcal{E} + \mathcal{E}_1 = 0. \tag{87-3}$$

If the current is varying, an electromotive force $\mathcal{E}_I$ (taken as positive when acting from a to b) is induced in the coil L on account of its self-inductance. This e.m.f. likewise gives rise to a separation of charge until a potential drop $\mathcal{E}_2$ is established between the ends a and b of the coil such that

$$\mathcal{E}_I + \mathcal{E}_2 = 0. \tag{87-4}$$

Finally the charges on the wire give rise to a difference of potential between the ends b and c of the resistance. Denoting

the drop of potential in passing from b to c by $\mathcal{E}_3$ we have

$$\mathcal{E}_3 = Ri \tag{87-5}$$

from Ohm's law.

Adding equations (87–3), (87–4) and (87–5),

$$\mathcal{E} + \mathcal{E}_I + (\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3) = Ri. \tag{87-6}$$

But the electromotive force taken all the way around the circuit due to the static charges vanishes, that is,

$$\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 = 0.$$

Hence, replacing $\mathcal{E}_I$ by the right-hand side of (87–2) and rearranging terms,

$$L\frac{di}{dt} + Ri = \mathcal{E}. \tag{87-7}$$

This relation is known as the *equation of the circuit*. It is valid in either e.m.u. or practical units.

It should be observed that the effect of the charges on the circuit is to transfer the place at which work is done on the electrons constituting the current from the region of the external electromotive force and the coil L to the resistance. As the resultant e.m.f. through the external source is zero by (87–3) no work is done there on the current, and the same is true of the coil L on account of (87–4). The entire work done on the current takes place in the resistance, the energy expended there being dissipated in the form of heat.

Nevertheless work is done by the external electromotive force in producing the separation of charge across ca and ab. This work gives rise to an increase in the energy of the circuit which is reversible in that it may be recovered as the current decreases. As will be shown later this energy can most naturally be associated with the magnetic field of the circuit.

If a voltmeter V (Fig. 185) is connected across the coil L of the circuit under consideration in such a way that there is no varying magnetic flux through the space between the leads, the

voltmeter measures the potential drop $\mathcal{E}_2$ between the terminals *a* and *b* of the coil. According to (87–4) this potential drop, which for convenience in later articles we shall designate by $\mathcal{E}_L$, is

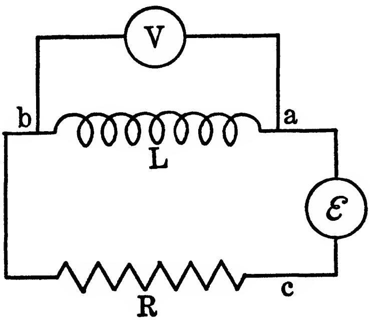


FIG. 185

$$\mathcal{E}_L = -\mathcal{E}_I = L\frac{di}{dt}.$$

The voltmeter reading, therefore, is equal to the negative of the electromotive force due to self-induction in the coil. If, for instance, the current is increasing, the e.m.f. of self-induction is directed from *b* to *a* whereas the voltmeter registers the equal potential drop from *a* to *b*. We shall refer to $\mathcal{E}_L$ as the electromotive force *across* the coil.

Heretofore we have supposed the resistance of the circuit to lie entirely outside the impressed e.m.f. and the inductance. Suppose now that a resistance R_1 is associated with the external e.m.f. and a resistance R_2 with the coil L, the remaining resistance of the circuit being denoted by R_3. Then equations (87–3), (87–4) and (87–5) become

$$\begin{aligned} \mathcal{E} + \mathcal{E}_1 &= R_1 i, \\ \mathcal{E}_I + \mathcal{E}_2 &= R_2 i, \\ \mathcal{E}_3 &= R_3 i, \end{aligned}$$

respectively, and adding we get

$$\mathcal{E} + \mathcal{E}_I + (\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3) = (R_1 + R_2 + R_3)i.$$

If we put R for the total resistance $R_1 + R_2 + R_3$ of the circuit and remember that $\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3$ is zero, we are led to (87–7) again. Therefore this equation is valid no matter how the resistance may be distributed in the circuit.

Growth and Decay of Current. — Suppose that the impressed electromotive force $\mathcal{E}$ is a constant e.m.f. such as might be supplied by a cell, and let the circuit be suddenly completed by

closing a key or switch. We wish to investigate the rate of growth of the current. Separating the variables i and t in (87–7),

$$\frac{di}{i - \dfrac{\mathcal{E}}{R}} = -\frac{R}{L}\,dt,$$

of which the integral is

$$i - \frac{\mathcal{E}}{R} = Ae^{-(R/L)t}.$$

If we determine the constant of integration A by taking $t = 0$ when $i = 0$,

$$i = \frac{\mathcal{E}}{R}\{1 - e^{-(R/L)t}\}. \qquad (87\text{–}8)$$

The current approaches asymptotically the value $\mathcal{E}/R$ given by Ohm's law (49–7) for the steady state. It reaches $(1 - 1/e)$th of its final value, that is, $0.632\ \mathcal{E}/R$, in the time L/R. This ratio,

$$\lambda = \frac{L}{R}, \qquad (87\text{–}9)$$

is called the *time constant* of the circuit. The smaller λ, the more rapidly the current grows. In Fig. 186 the current is plotted

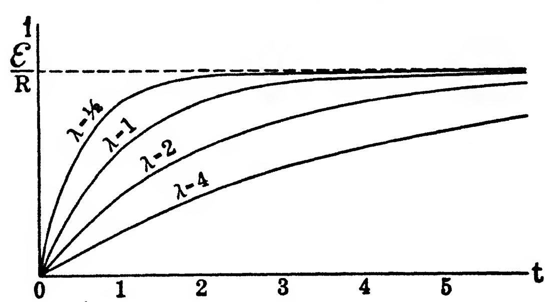


FIG. 186

against the time for several different values of the time constant. As t and λ each represent a time, the units employed on the diagram are arbitrary.

Next let us suppose that the impressed e.m.f. is suddenly removed from the circuit while a current i_0 is flowing, without the circuit being broken. We wish to find the law of decay of the current. In this case $\mathcal{E}$ vanishes and (87–7) becomes

$$\frac{di}{i} = -\frac{R}{L}dt,$$

leading to the integral

$$i = i_0 e^{-(R/L)t}, \tag{87–10}$$

if we count time from the instant at which the impressed e.m.f. is removed. The current decays exponentially, falling to $1/e$th of its original value, that is, 0.368 i_0, in a time λ. The decay curves for the values of λ appearing in Fig. 186 are obtained by inverting the figure. The smaller λ, the more rapidly the current falls off.

Simple Harmonic Electromotive Force. — If the impressed electromotive force $\mathcal{E}$ is a simple harmonic function of the time,

$$\mathcal{E} = \mathcal{E}_0 \sin \omega t,$$

such as might be produced by a coil rotating with constant angular velocity ω in a uniform magnetic field, the equation of the circuit becomes

$$L\frac{di}{dt} + Ri = \mathcal{E}_0 \sin \omega t. \tag{87–11}$$

Evidently in the steady state the current alternates with the same frequency as the impressed e.m.f. Therefore we should expect a solution of (87–11) of the form

$$i = i_0 \sin (\omega t - \phi).$$

Substituting in (87–11) and arranging terms,

$$(L\omega i_0 \cos \phi - Ri_0 \sin \phi) \cos \omega t + (L\omega i_0 \sin \phi + Ri_0 \cos \phi - \mathcal{E}_0) \sin \omega t = 0,$$

and the differential equation is satisfied for all values of t provided the coefficients of cos ωt and sin ωt vanish, that is,

$$\tan \phi = \frac{L\omega}{R}, \qquad i_0 = \frac{\mathcal{E}_0}{L\omega \sin \phi + R \cos \phi}.$$

Consequently, if we put the values of $\sin\phi$ and $\cos\phi$ obtained from the first of these relations in the second,

$$i = \frac{\mathcal{E}_0}{\sqrt{R^2 + L^2\omega^2}} \sin(\omega t - \phi), \qquad \tan\phi \equiv \frac{L\omega}{R}. \qquad (87\text{–}12)$$

The current, therefore, is out of phase with the impressed e.m.f., *lagging* behind it by an angle ϕ, that is, by $\phi/2\pi$ of a period. The tangent of the *lag* is proportional to the frequency $\omega/2\pi$ as well as to the self-inductance, the lag increasing from 0 for a steady e.m.f. to $\pi/2$ for infinite frequency.

The quantity

$$Z = \sqrt{R^2 + L^2\omega^2}$$

is known as the *impedance* of the circuit. The amplitude i_0 of the current is related to the amplitude $\mathcal{E}_0$ of the impressed e.m.f. by the equation

$$i_0 = \frac{\mathcal{E}_0}{\sqrt{R^2 + L^2\omega^2}}. \qquad (87\text{–}13)$$

Therefore impedance plays much the same part in this a.c. (alternating current) circuit that resistance does in a d.c. (direct current) circuit. It is to be noted that the impedance increases with the frequency. The quantity $L\omega$ appearing in the expression for the impedance is called the *reactance*. At low frequency the reactance may be relatively unimportant as compared with the resistance, whereas at high frequency the reactance may be so great as to make the effect of the resistance quite negligible. The ohm is the practical unit of impedance and of reactance as well as of resistance.

Equation (87–12) is a particular solution of the equation of the circuit in that it contains no arbitrary constant. As (87–11) is a differential equation of the first order, the complete solution must contain one constant of integration. For the moment let us denote the particular solution (87–12) by i_2 and the solution of the equation obtained by making the right-hand member of (87–11) equal to zero by i_1. Then

$$i = i_1 + i_2$$

is a solution of (87–11). For, if we substitute in the differential equation,

$$\left(L\frac{di_1}{dt} + Ri_1\right) + \left(L\frac{di_2}{dt} + Ri_2 - \mathcal{E}_0 \sin \omega t\right) = 0,$$

which is satisfied identically since each of the two expressions within parentheses vanishes. Now i_1 is the solution (87–10) already obtained. So if we replace the coefficient i_0 in (87–10) by A to indicate that it represents an arbitrary constant of integration,

$$i = Ae^{-(R/L)t} + \frac{\mathcal{E}_0}{\sqrt{R^2 + L^2\omega^2}} \sin(\omega t - \phi), \quad \tan\phi \equiv \frac{L\omega}{R}, \quad (87\text{–}14)$$

is the complete solution of the equation of the circuit. The constant A is to be determined by the initial conditions, for instance by the current at the time $t = 0$. The current represented by the first term is known as a *transient*. On account of the negative sign in the exponent the magnitude of this term

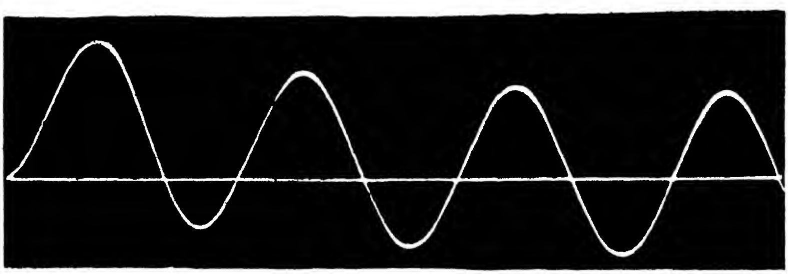


Fig. 187

decreases exponentially with the time, so that after a sufficiently long interval it becomes negligible as compared with the second term, and the current becomes that of the steady state represented by (87–12). The oscillogram (art. 115) shown in Fig. 187, which represents the current from the instant of connecting the circuit with a simple harmonic e.m.f., illustrates the decay of the transient and the establishment of a sinusoidal current.

Problem 87a. A circuit consists of two concentric cylindrical shells of radii a and b ($b > a$) and common length l connected by flat end plates, the current flowing out along one shell and back along the other. The space between the shells contains a medium of permeability μ. Find the self-inductance of the circuit in e.m.u.

$$\text{Ans. } 2\mu l \log_e \frac{b}{a}\cdot$$

Problem 87b. Find the self-inductance in e.m.u. of a toroidal solenoid of mean radius a and cross-sectional radius b containing n turns of wire wound on a ring of permeability μ.

$$\text{Ans. } 4\pi\mu n^2(a - \sqrt{a^2 - b^2}).$$

Problem 87c. A circuit contains a ring solenoid of 20 cm radius and 5 cm² cross-section consisting of 2000 turns. It is wound on an iron core of permeability 2000 and has a resistance of 10 ohms. Find the time constant λ. Ans. 0.4 sec.

Problem 87d. A circuit has a resistance of 2 ohms and a self-inductance of 0.01 henry. At what frequency is the reactance equal to the resistance? What is the lag of the current at this frequency? Ans. 31.8 cycle, $\pi/4$.

Problem 87e. In the circuit of 87*c*, how long will it take a transient to fall to 1% of its initial value? Ans. 1.8 sec.

Problem 87f. If a part R_2 of the resistance of the circuit discussed in this article is associated with the coil L, what will be the reading on a voltmeter connected across the coil? Ans. $\mathcal{E}_L + R_2 i$.

Problem 87g. A simple a.c. generator such as illustrated in Fig. 182 is connected to an external resistance of 1 ohm. The armature coil has a self-inductance of 2 millihenrys and rotates 60 times a second. It has an area of 100 cm² and contains 100 turns. The magnetic field in which it rotates is 1000 gauss. Find the current. Ans. $i = 30.2 \sin(\omega t - 0.65)$ amp.

Problem 87h. A circuit contains a constant external e.m.f. and an inductance but no resistance. According to what law does the current grow? Ans. $i = \dfrac{\mathcal{E}}{L}t$.

88. Circuit with Capacity and Resistance. — Consider a condenser C (Fig. 188) in series with an impressed electromotive force $\mathcal{E}$ and a resistance R, the self-inductance of the circuit being so small as to be negligible. If $\mathcal{E}$ is a constant e.m.f. acting from c to a, the plate a of the condenser will acquire a positive charge and the plate b an equal negative charge of magnitude sufficient to produce a difference of potential between the plates equal to $\mathcal{E}$,

the current ceasing to flow as soon as this state has been attained. On the other hand, if $\mathcal{E}$ is an alternating e.m.f., current will flow back and forth through the circuit from the one to the other plate of the condenser, the charge on each plate reversing sign with every alternation of the electromotive force.

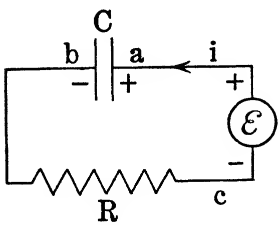


FIG. 188

Let us analyze the circuit. If the resistance associated with $\mathcal{E}$ is negligible the charges present must give rise to a drop of potential $\mathcal{E}_1$ between c and a just sufficient to neutralize $\mathcal{E}$, so that at every instant

$$\mathcal{E} + \mathcal{E}_1 = 0. \tag{88–1}$$

Next there is a potential drop $\mathcal{E}_2$ across the condenser due to the charges q and $-q$ on the plates. If C denotes its capacity measured in the same system of units as $\mathcal{E}_2$ and q,

$$\mathcal{E}_2 = \frac{q}{C}. \tag{88–2}$$

Finally the charges present on the circuit give rise to a difference of potential $\mathcal{E}_3$ between the ends b and c of the resistance amounting to

$$\mathcal{E}_3 = Ri. \tag{88–3}$$

Adding (88–1), (88–2) and (88–3),

$$\mathcal{E} + (\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3) = \frac{q}{C} + Ri.$$

But the electromotive force $(\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3)$ around the entire circuit due to charges on the condenser plates and on the wire vanishes. Hence, as $i = \dfrac{dq}{dt}$,

$$R\frac{dq}{dt} + \frac{q}{C} = \mathcal{E} \tag{88–4}$$

is the equation of the circuit under consideration. This equation is identical in form with the equation (87–7) of a circuit containing self-inductance and resistance only, q taking the place of i, R that of L and $1/C$ that of R.

Growth and Decay of Charge. — Suppose that $\mathcal{E}$ is a constant e.m.f. and that the circuit has just been completed by closing a key, the condenser having previously been discharged. We wish to investigate the rate at which charge accumulates on the condenser plates. As the equation of the circuit is of the same form as (87–7), we have at once from (87–8)

$$q = C\mathcal{E}\{1 - e^{-(1/RC)t}\}, \qquad (88\text{–}5)$$

by substituting $1/C$ for R and R for L. The charge increases exponentially to the final value $C\mathcal{E}$. The time constant,

$$\lambda = RC, \qquad (88\text{–}6)$$

represents the time necessary for the charge to attain $(1 - 1/e)$th of its final value. The smaller λ the more rapidly the charge accumulates. If we interpret the ordinates in Fig. 186 as charge on the condenser instead of current in the circuit, the curves there drawn show the accumulation of charge for various values of the time constant.

To get the current in the circuit (88–5) must be differentiated with respect to the time, giving

$$i = \frac{dq}{dt} = \frac{\mathcal{E}}{R} e^{-(1/RC)t}. \qquad (88\text{–}7)$$

The current starts with the value $\mathcal{E}/R$, gradually decreasing and finally stopping when the condenser has attained its full charge $C\mathcal{E}$.

Next consider a circuit consisting of a condenser with charge q_0, key and resistance connected in series, no impressed e.m.f. being present. We wish to find the way in which the condenser discharges through the resistance after the key is closed. All we have to do is to make the appropriate substitutions in (87–10), getting

$$q = q_0 e^{-(1/RC)t}. \qquad (88\text{–}8)$$

We note that the condenser discharges exponentially, the rate of discharge or the current being

$$i = \frac{dq}{dt} = -\frac{q_0}{RC}e^{-(1/RC)t}, \tag{88-9}$$

where the negative sign indicates that the current is in the sense to decrease the charge on the condenser. The discharge is more rapid the smaller the time constant.

Simple Harmonic Electromotive Force.—Now we shall suppose that the impressed electromotive force $\mathcal{E}$ in the circuit of Fig. 188 is a simple harmonic function of the time of the form

$$\mathcal{E} = \mathcal{E}_0 \sin \omega t.$$

Then the equation of the circuit is

$$R\frac{dq}{dt} + \frac{q}{C} = \mathcal{E}_0 \sin \omega t. \tag{88-10}$$

As this equation is of the same form as (87–11) we obtain the solution for the steady state from (87–12) by means of the substitutions previously employed. Thus,

$$q = \frac{\mathcal{E}_0}{\sqrt{\frac{1}{C^2} + R^2\omega^2}} \sin (\omega t - \psi),$$

$$= \frac{\frac{\mathcal{E}_0}{\omega}}{\sqrt{R^2 + \frac{1}{C^2\omega^2}}} \sin (\omega t - \psi), \qquad \tan \psi \equiv RC\omega, \tag{88-11}$$

and the current is

$$i = \frac{dq}{dt} = \frac{\mathcal{E}_0}{\sqrt{R^2 + \frac{1}{C^2\omega^2}}} \cos (\omega t - \psi)$$

$$= \frac{\mathcal{E}_0}{\sqrt{R^2 + \frac{1}{C^2\omega^2}}} \sin (\omega t - \phi), \qquad \tan \phi \equiv -\frac{1}{RC\omega}. \tag{88-12}$$

As in the case of a circuit containing self-inductance and

resistance the current is out of phase with the impressed e.m.f., but here the current *leads* the e.m.f. instead of lagging behind it since ϕ is negative. For high frequency the angle ϕ is very nearly zero, decreasing to $-\pi/2$ for zero frequency. Note that this is quite different from the case of a circuit containing self-inductance.

The *impedance* of the circuit is

$$Z = \sqrt{R^2 + \frac{1}{C^2\omega^2}},$$

the amplitude i_0 of the current being related to that of the e.m.f. by the equation

$$i_0 = \frac{\mathcal{E}_0}{\sqrt{R^2 + \frac{1}{C^2\omega^2}}}. \tag{88–13}$$

Contrary to what was found in the case of self-inductance, the impedance of the circuit under consideration here decreases as the frequency is increased, approaching the limiting value R for very high frequencies. For zero frequency, that is a steady e.m.f., the impedance is infinite and no current can flow. Obviously this must be true, for the condenser constitutes a break in the circuit and no steady current can flow under the action of a constant e.m.f.

The *reactance* of the circuit is $-1/C\omega$. It is important only at low frequencies, becoming negligible at high enough frequencies. In fact the insertion of a condenser in a circuit has no appreciable effect upon either the amplitude or the phase of the current at sufficiently high frequency.

So far we have omitted the transient. The complete solution of the equation of the circuit may be obtained from (87–14) by means of the substitutions already employed or by adding (88–9) to (88–12). It is

$$i = Ae^{-(1/RC)t} + \frac{\mathcal{E}_0}{\sqrt{R^2 + \frac{1}{C^2\omega^2}}}\sin(\omega t - \phi), \qquad \tan\phi \equiv -\frac{1}{RC\omega}, \tag{88–14}$$

the constant of integration A being determined from the initial value of the current. The transient current, represented by the first term, falls off with the time, leaving the second term alone after the steady state has been attained.

Problem 88a. A condenser of 2 microfarads capacity is connected in series with a high resistance and a quadrant electrometer. It is found that the condenser loses half its charge in 120 sec. Find the resistance. (Measurement of high resistance by leakage.) Ans. 86.6 megohm.

Problem 88b. A leaky condenser of capacity C and conductivity K is connected in series with a resistance R and an impressed e.m.f. $\mathcal{E}_0 \sin \omega t$. Find the impedance of the circuit.

$$\text{Ans. } \sqrt{R^2 + \frac{1 + 2KR}{C^2\omega^2 + K^2}}.$$

89. Circuit with Self-Inductance, Capacity and Resistance. — A more general type of circuit than those considered heretofore is one in which a self-inductance L (Fig. 189), a condenser of capacity C and a resistance R are connected in series with an impressed electromotive force $\mathcal{E}$. Denoting the potential drop from d to a due to the static charges on the circuit by $\mathcal{E}_1$, that from a to b by $\mathcal{E}_2$, that between the plates b and c of the condenser by $\mathcal{E}_3$, and that across the resistance by $\mathcal{E}_4$,

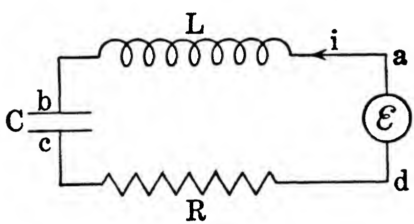


FIG. 189

$$\mathcal{E} + \mathcal{E}_1 = 0,$$

$$\mathcal{E}_I + \mathcal{E}_2 = 0,$$

$$\mathcal{E}_3 = \frac{q}{C},$$

$$\mathcal{E}_4 = Ri,$$

where $\mathcal{E}_I$ denotes, as before, the induced e.m.f.

$$\mathcal{E}_I = -L\frac{di}{dt} \tag{89–1}$$

due to the inductance of the coil L. Adding the preceding equations,

$$\mathcal{E} + \mathcal{E}_I + (\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 + \mathcal{E}_4) = Ri + \frac{q}{C}.$$

As the e.m.f. around the complete circuit due to the static charges vanishes, the expression within parentheses is zero. So, making use of (89–1) and rearranging terms,

$$L\frac{di}{dt} + Ri + \frac{q}{C} = \mathcal{E}. \tag{89–2}$$

To obtain the differential equation for q we put $\dfrac{dq}{dt}$ for i getting

$$L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{q}{C} = \mathcal{E}, \tag{89–3}$$

or, if we wish to use i as dependent variable, we differentiate (89–2) with respect to the time, so as to have the equation of the circuit in the form

$$L\frac{d^2i}{dt^2} + R\frac{di}{dt} + \frac{1}{C}i = \frac{d\mathcal{E}}{dt}. \tag{89–4}$$

Equations (89–3) and (89–4) hold for either e.m.u. or practical units. As the reader can easily show they are equally valid when part of the resistance is associated with the coil L and part with the electromotive force $\mathcal{E}$. Furthermore part of the self-inductance may be associated with $\mathcal{E}$, and there may be several condensers in the circuit. In every case, however, R represents the total resistance in the circuit, L the total self-inductance and C the total capacity.

Charging Condenser. — If the e.m.f. in the circuit under consideration is constant, we can find the rate at which charge accumulates on the condenser plates when the circuit is completed by integrating (89–3), the right-hand member being treated as a constant. The equation is put in more convenient form if we introduce a new dependent variable,

$$Q = q - C\mathcal{E}.$$

Then (89–3) becomes

$$L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{Q}{C} = 0. \tag{89–5}$$

As this equation is linear and homogeneous in Q, the solution is of the form

$$Q = Ae^{\gamma t},$$

where A is an arbitrary constant. Substituting in (89–5),

$$L\gamma^2 + R\gamma + \frac{1}{C} = 0,$$

giving

$$\gamma = -\frac{R}{2L} \pm \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}}. \tag{89–6}$$

Each value of γ gives a solution. If we put

$$\alpha \equiv \frac{R}{2L}, \qquad \beta \equiv \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}},$$

so that $\gamma = -\alpha \pm \beta$, we have for the complete solution of (89–5)

$$q = C\mathcal{E} + e^{-\alpha t}\{A_1e^{\beta t} + A_2e^{-\beta t}\}, \tag{89–7}$$

where A_1 and A_2 are constants to be determined by the initial conditions. The current is obtained by differentiating (89–7) with respect to the time:

$$i = e^{-\alpha t}\{(\beta - \alpha)A_1e^{\beta t} - (\beta + \alpha)A_2e^{-\beta t}\}. \tag{89–8}$$

If the condenser is uncharged when the key is closed, $q = i = 0$ when $t = 0$. Consequently

$$A_1 + A_2 = -C\mathcal{E}$$

from (89–7), and

$$(\beta - \alpha)A_1 - (\beta + \alpha)A_2 = 0$$

from (89–8). Solving for A_1 and A_2,

$$A_1 = -\frac{\alpha + \beta}{2\beta}C\mathcal{E}, \qquad A_2 = \frac{\alpha - \beta}{2\beta}C\mathcal{E},$$

and

$$q = C\mathcal{E}\left[1 - \frac{1}{2\beta}e^{-\alpha t}\{(\alpha + \beta)e^{\beta t} - (\alpha - \beta)e^{-\beta t}\}\right]. \qquad (89\text{–}9)$$

The constant β is real if

$$\frac{R^2}{4L^2} > \frac{1}{LC},$$

that is, if the resistance is larger than twice the square root of the ratio of self-inductance to capacity. In this case $\beta < \alpha$ by (89–6), the term involving the exponentials decreases continuously with the time, and q approaches the limiting value $C\mathcal{E}$. The broken curve in Fig. 190 shows how q increases with the time. Evidently the charging of the condenser is aperiodic.

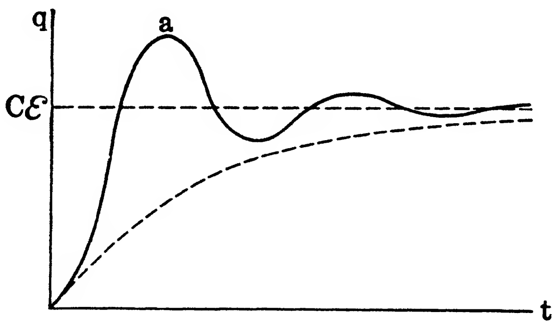


FIG. 190

If the resistance of the circuit is less than twice the square root of the ratio of self-inductance to capacity β becomes imaginary and the solution (89–9) is not in convenient form. In this case oscillations are set up in the circuit, and we may proceed by assuming a solution of (89–5) of the form

$$Q = Ae^{-\alpha t}\sin(\omega_0 t + \epsilon). \qquad (89\text{–}10)$$

Substituting in the differential equation,

$$\left\{(\alpha^2 - \omega_0^2)L - \alpha R + \frac{1}{C}\right\}\sin(\omega_0 t + \epsilon)$$
$$+ \{-2\alpha\omega_0 L + \omega_0 R\}\cos(\omega_0 t + \epsilon) = 0,$$

which is satisfied for all values of t provided the coefficients of the sine and cosine vanish, that is,

$$\alpha = \frac{R}{2L}, \qquad \omega_0 = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}. \tag{89-11}$$

Consequently the complete solution of (89–5) is

$$q = C\mathcal{E} + Ae^{-\alpha t} \sin (\omega_0 t + \epsilon), \tag{89-12}$$

A and ϵ being arbitrary constants to be determined by the initial conditions. As α is positive, the amplitude of the oscillations given by the second term falls off with the time and q approaches the same final value $C\mathcal{E}$ as in the previous case. The current,

$$i = Ae^{-\alpha t}\{\omega_0 \cos (\omega_0 t + \epsilon) - \alpha \sin (\omega_0 t + \epsilon)\}, \tag{89-13}$$

is also oscillatory, becoming, however, of smaller and smaller amplitude as time elapses.

Instead of solving the differential equation of the circuit anew for the oscillatory case we may put

$$\beta \equiv \mathrm{i}\omega_0 \equiv \mathrm{i}\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}$$

in (89–7), where $\mathrm{i} \equiv \sqrt{-1}$, and obtain (89–12) immediately by means of the identity

$$e^{\pm \mathrm{i}\omega_0 t} \equiv \cos \omega_0 t \pm \mathrm{i} \sin \omega_0 t.$$

For then the expression inside the braces in the second term of (89–7) becomes

$$(A_1 + A_2) \cos \omega_0 t + \mathrm{i}(A_1 - A_2) \sin \omega_0 t,$$

which reduces to

$$A \sin (\omega_0 t + \epsilon),$$

if we put

$$A \equiv 2\sqrt{A_1 A_2}, \qquad \tan \epsilon \equiv \mathrm{i}\frac{A_2 + A_1}{A_2 - A_1}.$$

Under the same initial conditions as considered previously, $q = i = 0$ when $t = 0$ in (89–12). Hence

$$A \sin \epsilon = - C\mathcal{E},$$
$$\omega_0 \cos \epsilon - \alpha \sin \epsilon = 0,$$

giving

$$\tan \epsilon = \frac{\omega_0}{\alpha}, \qquad A = -\frac{C\mathcal{E}}{\sin \epsilon} = -\frac{\sqrt{\omega_0^2 + \alpha^2}}{\omega_0} C\mathcal{E} \cdot$$

Therefore

$$q = C\mathcal{E}\left\{1 - \frac{\sqrt{\omega_0^2 + \alpha^2}}{\omega_0} e^{-\alpha t} \sin (\omega_0 t + \epsilon)\right\},$$

$$\tan \epsilon \equiv \frac{\omega_0}{\alpha}; \qquad (89\text{–}14)$$

and the current is

$$i = \frac{\sqrt{\omega_0^2 + \alpha^2}}{\omega_0} C\mathcal{E}e^{-\alpha t}\{\alpha \sin (\omega_0 t + \epsilon) - \omega_0 \cos (\omega_0 t + \epsilon)\}$$

$$= \frac{\omega_0^2 + \alpha^2}{\omega_0} C\mathcal{E}e^{-\alpha t} \sin \omega_0 t,$$

which we may write

$$i = i_0 e^{-\alpha t} \sin \omega_0 t, \qquad i_0 \equiv \frac{\omega_0^2 + \alpha^2}{\omega_0} C\mathcal{E}. \qquad (89\text{–}15)$$

The charge on the condenser in this case is shown by the solid curve in Fig. 190. During the oscillations the charge may become considerably greater than the final value $C\mathcal{E}$, and consequently the potential difference between the plates may exceed $\mathcal{E}$, as at the point a on the curve. This sometimes results in a breakdown of the dielectric between the plates, entailing destruction of the condenser. Therefore care must be exercised in connecting a circuit containing a condenser and a self-inductance in series to a source of potential in order to avoid excessive oscillations. This may be done by including in the circuit sufficient resistance. If the resistance is great enough, the charge increases aperiodically in conformity with the broken curve, and never becomes greater than $C\mathcal{E}$.

The factor $e^{-\alpha t}$ in the equation (89–15) causes the amplitude of the current oscillations to become smaller and smaller as time goes on. This factor is known as a *damping factor*. The larger α the more rapidly the current is damped out.

Discharging Condenser. — If there is no impressed e.m.f. in the circuit, (89–3) becomes identical with (89–5), Q representing the charge on the condenser. This is the case of a charged condenser connected in series with an inductance, resistance and key. The solution of (89–5) specifies the rate of discharge of the condenser through the inductance and resistance after the key is closed.

If the resistance is large the discharge is aperiodic, the charge q being obtained from (89–7) by making $\mathcal{E}$ zero:

$$q = e^{-\alpha t}\{A_1 e^{\beta t} + A_2 e^{-\beta t}\}. \tag{89–16}$$

The current is still given by (89–8). To determine A_1 and A_2 we have $q = q_0$ and $i = 0$ when $t = 0$. Therefore

$$A_1 + A_2 = q_0,$$
$$(\beta - \alpha)A_1 - (\beta + \alpha)A_2 = 0,$$

giving

$$A_1 = \frac{\alpha + \beta}{2\beta} q_0, \qquad A_2 = -\frac{\alpha - \beta}{2\beta} q_0.$$

Consequently (89–16) becomes

$$q = \frac{q_0}{2\beta} e^{-\alpha t}\{(\alpha + \beta)e^{\beta t} - (\alpha - \beta)e^{-\beta t}\}. \tag{89–17}$$

The discharge curve is represented by the broken line in Fig. 190 provided the figure is inverted.

When the resistance is so small as to make β imaginary, we need the periodic solution obtained from (89–12) by making $\mathcal{E}$ zero. In this case

$$q = Ae^{-\alpha t}\sin(\omega_0 t + \epsilon), \tag{89–18}$$

the current still being given by (89–13). Making use of the initial conditions employed above,

$$A \sin\epsilon = q_0,$$
$$\omega_0 \cos\epsilon - \alpha \sin\epsilon = 0,$$

and

$$\tan\epsilon = \frac{\omega_0}{\alpha}, \qquad A = \frac{q_0}{\sin\epsilon} = \frac{\sqrt{\omega_0^2 + \alpha^2}}{\omega_0} q_0.$$

Consequently

$$q = \frac{\sqrt{\omega_0^2 + \alpha^2}}{\omega_0} q_0 e^{-\alpha t} \sin (\omega_0 t + \epsilon), \tag{89–19}$$

and the current is

$$\begin{aligned} i &= \frac{\sqrt{\omega_0^2 + \alpha^2}}{\omega_0} q_0 e^{-\alpha t} \{\omega_0 \cos (\omega_0 t + \epsilon) - \alpha \sin (\omega_0 t + \epsilon)\} \\ &= -\frac{\omega_0^2 + \alpha^2}{\omega_0} q_0 e^{-\alpha t} \sin \omega_0 t, \end{aligned}$$

which we will write in the more convenient form

$$i = i_0 e^{-\alpha t} \sin \omega_0 t, \qquad i_0 \equiv -\frac{\omega_0^2 + \alpha^2}{\omega_0} q_0. \tag{89–20}$$

The current is plotted against the time in Fig. 191, the dotted

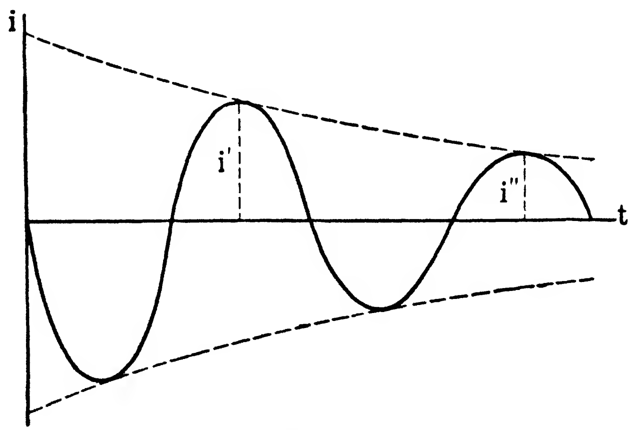


Fig. 191

lines representing the enveloping curves

$$i = \pm\, i_0 e^{-\alpha t}.$$

The *natural frequency* of the circuit is

$$\nu_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi}\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}, \tag{89–21}$$

the *natural* or *free period* P_0 being the reciprocal of this. If R is small compared to $2\sqrt{L/C}$, which is often the case,

$$\nu_0 = \frac{\omega_0}{2\pi} = \frac{1}{2\pi\sqrt{LC}} \tag{89–22}$$

to a sufficient degree of approximation. In this case the inductive reactance $L\omega_0$ of the circuit is equal to the magnitude $1/C\omega_0$ of the capacitative reactance.

To find the time t_m when the current is maximum we place the derivative of (89–19) with respect to t equal to zero, getting

$$\tan \omega_0 t_m = \frac{\omega_0}{\alpha} \cdot$$

If T_0 is the time elapsing between successive maxima,

$$\omega_0(t_m + T_0) = \omega_0 t_m + 2\pi,$$

or

$$T_0 = \frac{2\pi}{\omega_0} = P_0. \tag{89–23}$$

The maxima recur, then, at intervals equal to the free period of oscillation. If i' and i'' (Fig. 191) are two successive maxima,

$$\frac{i'}{i''} = \frac{e^{-\alpha t}}{e^{-\alpha(t+P_0)}} = e^{\alpha P_0}.$$

The quantity δ defined by

$$\delta = \log \frac{i'}{i''} = \alpha P_0 = \frac{2\pi\left(\dfrac{R}{2L}\right)}{\sqrt{\dfrac{1}{LC} - \dfrac{R^2}{4L^2}}} \tag{89–24}$$

is known as the *logarithmic decrement* of the circuit. If the resistance of the circuit is small so that (89–22) holds,

$$\delta = \pi R\sqrt{\frac{C}{L}} = \pi\frac{R}{L\omega_0} = \pi\frac{R}{\dfrac{1}{C\omega_0}}, \tag{89–25}$$

showing that the logarithmic decrement is equal to π times the ratio of the resistance to the magnitude of either the inductive or

the capacitative reactance. The logarithmic decrement is a convenient index of the rate at which the oscillations are damped out.

The oscillator used by Hertz to produce electromagnetic waves in his famous experiments of 1888 consisted essentially of a condenser discharging through an inductance and resistance. The apparatus is shown in Fig. 192. The two metal plates A and B form the condenser, the self-inductance of the circuit being that of the two short wires connecting the plates to the spark gap G. As the first discharge ionizes the air in the gap, the resistance becomes so small that oscillations are set up. The potential difference existing between A and B in the first place may be produced by connecting the two sides of the gap to the secondary of an induction coil. Each time the primary current in the coil is broken, an oscillatory discharge passes across the gap, as illustrated in Fig. 193 for two breaks of the primary. The impedance of the secondary of the coil is so great for the high frequency discharge taking place across the gap that no appreciable part of the current passes through the coil. As will be shown in Chapter XVI an oscillatory discharge such as that described here acts as a source of electromagnetic waves which travel out from it through the surrounding space.

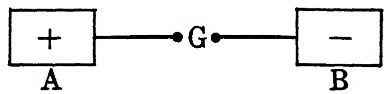


FIG. 192

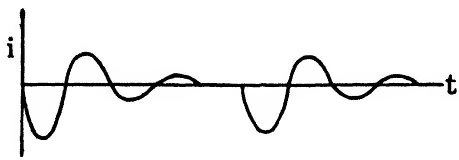


FIG. 193

Problem 89a. A circuit has an inductance of 10 millihenrys, capacity of 0.1 microfarad and resistance of 1000 ohms. Is it periodic or not? If there is an impressed e.m.f. of $(10)^5$ volts in the circuit, what is the final charge on the condenser? Ans. 0.01 coulomb.

Problem 89b. A condenser of capacity 1 microfarad is connected in series with an inductance of 10 millihenrys and a resistance of 10 ohms. Initially the condenser has a charge of 1 coulomb. Find the natural frequency of the circuit and the charge at the end of 10 oscillations. Ans. 1590 cycle, 0.231 coulomb.

Problem 89c. An oscillatory circuit is connected with a constant e.m.f. Show that the greatest charge acquired by the condenser is

$$q_m = C\mathcal{E}\{1 + e^{-(\pi\alpha/\omega_0)}\}.$$

If C, L and R have the values given in 89*b*, find the ratio of q_m to the final charge. Ans. 1.86.

Problem 89d. A circuit in which

$$\frac{R^2}{4L^2} = \frac{1}{LC}$$

is said to be *critically damped*. Show that in this case a condenser discharges aperiodically according to the law,

$$q = q_0(1 + \alpha t)e^{-\alpha t}.$$

Problem 89e. Find the logarithmic decrement of a circuit containing an inductance of 10 millihenrys, a condenser of 1 microfarad capacity and a resistance of 1 ohm. By what percent does the current diminish per oscillation on discharge? Ans. 0.0314, 3.09%.

90. Alternating Current. — If the circuit discussed in the last article is connected with a simple harmonic impressed electromotive force

$$\mathcal{E} = \mathcal{E}_0 \sin \omega t,$$

the equation (89–4) of the circuit becomes

$$L\frac{d^2i}{dt^2} + R\frac{di}{dt} + \frac{1}{C}i = \mathcal{E}_0\omega \cos \omega t. \qquad (90\text{–}1)$$

In the steady state the current must alternate with the frequency of the impressed e.m.f. Therefore we look for a solution of the form

$$i = i_0 \sin (\omega t - \phi).$$

Substituting in (90–1) and arranging terms,

$$\left\{-L\omega^2 \cos \phi + R\omega \sin \phi + \frac{1}{C}\cos \phi\right\} i_0 \sin \omega t$$
$$+ \left\{\left(L\omega^2 \sin \phi + R\omega \cos \phi - \frac{1}{C}\sin \phi\right) i_0 - \mathcal{E}_0\omega\right\} \cos \omega t = 0,$$

showing that to satisfy the differential equation we must have

$$\tan \phi = \frac{L\omega - \dfrac{1}{C\omega}}{R}, \quad i_0 = \frac{\mathcal{E}_0}{R \cos \phi + \left(L\omega - \dfrac{1}{C\omega}\right) \sin \phi}.$$

Putting in the expression for i_0 the values of $\sin\phi$ and $\cos\phi$ obtained from the expression for $\tan\phi$,

$$i = \frac{\mathcal{E}_0}{\sqrt{R^2+\left(L\omega - \frac{1}{C\omega}\right)^2}}\sin(\omega t - \phi), \quad \tan\phi \equiv \frac{L\omega - \frac{1}{C\omega}}{R}; \qquad (90\text{–}2)$$

the relation between the amplitude of the current and that of the impressed e.m.f. being

$$i_0 = \frac{\mathcal{E}_0}{\sqrt{R^2+\left(L\omega - \frac{1}{C\omega}\right)^2}}. \qquad (90\text{–}3)$$

At low frequency the dominating term in the impedance,

$$Z = \sqrt{R^2+\left(L\omega - \frac{1}{C\omega}\right)^2}, \qquad (90\text{–}4)$$

is the one involving capacity, whereas at high frequency the term containing the inductance is the important one. The impedance is a minimum and therefore the current greatest for the frequency ν_r which makes the reactance $L\omega - 1/C\omega$ vanish, that is, for

$$\nu_r = \frac{\omega_r}{2\pi} = \frac{1}{2\pi\sqrt{LC}}. \qquad (90\text{–}5)$$

By analogy with the simple harmonic oscillator this frequency is known as the *frequency of resonance*. In a circuit of negligible

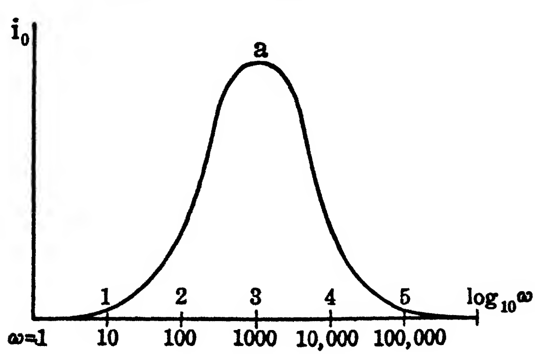


FIG. 194

resistance it is equal to the natural frequency (89–22). In Fig. 194 the amplitude of the current is plotted against the logarithm of the frequency for constant $\mathcal{E}_0$, the point a on the curve corresponding to the state of resonance. If we denote the ratio ω/ω_r by Ω equation (90–4) becomes

$$Z = \sqrt{R^2 + \frac{L}{C}\left(\Omega - \frac{1}{\Omega}\right)^2},$$

showing that the impedance is the same for the frequency $k\omega_r$ as for ω_r/k. Therefore the curve of Fig. 194 is symmetrical about the frequency of resonance.

For frequencies less than resonance ϕ is negative, indicating that the current leads the impressed e.m.f., the difference in phase, however, always being less than $\pi/2$. At resonance $\phi = 0$ and the current is in phase with the electromotive force. For frequencies greater than resonance ϕ is positive, the current lagging behind the e.m.f. by an angle between 0 and $\pi/2$. In terms of $\Omega = \omega/\omega_r$,

$$\tan\phi = \frac{\sqrt{\frac{L}{C}}\left(\Omega - \frac{1}{\Omega}\right)}{R},$$

showing that ϕ has the same absolute value (although opposite sign) for the frequency $k\omega_r$ as for ω_r/k. Therefore, if we plot ϕ

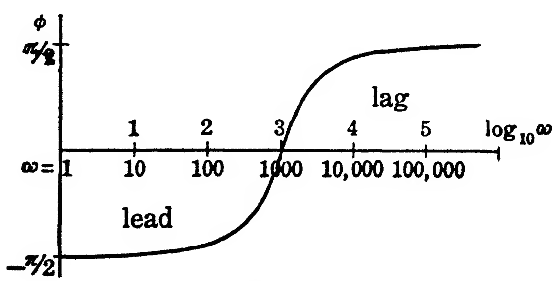


FIG. 195

against the logarithm of ω we get the symmetrical curve shown in Fig. 195.

The power delivered to the circuit by the impressed e.m.f. at

any instant t is

$$\mathcal{P} = \mathcal{E}i = \frac{\mathcal{E}_0^2}{\sqrt{R^2 + \left(L\omega - \frac{1}{C\omega}\right)^2}} \sin \omega t \sin (\omega t - \phi). \qquad (90\text{-}6)$$

Evidently the power fluctuates during the course of an oscillation, and, as the current is not in phase with the e.m.f., it even becomes negative for part of the time. If we represent the electromotive force by a solid curve (Fig. 196) and the current by a broken curve we can find the power at any instant by taking the product of the ordinates of the two curves, the power being

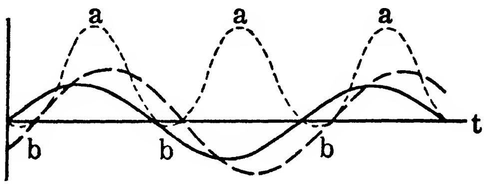


FIG. 196

positive when the ordinates have the same sign and negative when they have opposite signs. In this way we get the dotted curve on the diagram. At a, a, $\cdots$ the power is positive, that is, the circuit is drawing energy from the source of electromotive force, whereas at b, b, $\cdots$ the power is negative, the circuit returning energy to the external source. As ϕ is always less than $\pi/2$ in absolute value the mean power, averaged over a period, is always positive. It is

$$\begin{aligned}\overline{\mathcal{P}} &= \frac{\omega}{2\pi}\int_0^{2\pi/\omega} \mathcal{E}i\,dt \\ &= \frac{\omega}{2\pi}\frac{\mathcal{E}_0^2}{\sqrt{R^2 + \left(L\omega - \frac{1}{C\omega}\right)^2}}\left\{\cos\phi \int_0^{2\pi/\omega} \sin^2 \omega t\,dt \right. \\ &\qquad\qquad\left. - \sin\phi\int_0^{2\pi/\omega} \sin\omega t \cos\omega t\,dt\right\} \\ &= \frac{1}{2}\frac{\mathcal{E}_0^2}{\sqrt{R^2 + \left(L\omega - \frac{1}{C\omega}\right)^2}}\cos\phi. \qquad (90\text{-}7)\end{aligned}$$

The *effective current* i_e is the steady current which will develop the same amount of heat in the resistance R as the alternating current actually flowing. We have from (49–8) for the energy converted into heat per unit time

$$Ri_e^2 = R\left\{\frac{\omega}{2\pi}\int_0^{2\pi/\omega} i^2 dt\right\} = Ri_0^2\left\{\frac{\omega}{2\pi}\int_0^{2\pi/\omega} \sin^2(\omega t - \phi)dt\right\}$$
$$= \tfrac{1}{2}Ri_0^2. \qquad (90\text{–}8)$$

Consequently

$$\frac{1}{\sqrt{2}}.$$

As i_e is the square root of the mean value of the square of i it is also known as the *root-mean-square* (r.m.s.) value of the current.

We define the *effective* or *root-mean-square* (r.m.s.) *electromotive force* $\mathcal{E}_e$ in analogous manner by

$$\mathcal{E}_e^2 = \frac{\omega}{2\pi}\int_0^{2\pi/\omega} \mathcal{E}^2 dt = \mathcal{E}_0^2\left\{\frac{\omega}{2\pi}\int_0^{2\pi/\omega} \sin^2 \omega t dt\right\}$$

giving

$$\mathcal{E}_e = \frac{1}{\sqrt{2}}\mathcal{E}_0.$$

Hence, from (90–3) and (90–4),

$$i_e = \frac{\mathcal{E}_e}{Z},$$

and from (90–7),

$$\overline{\mathcal{P}} = \mathcal{E}_e i_e \cos\phi. \qquad (90\text{–}10)$$

The quantity $\cos\phi$ is known as the *power factor*. At resonance it is unity and the maximum of power is delivered to the circuit. At frequencies far from resonance the power factor becomes very small, and nearly as much work is done against the impressed e.m.f. during the course of an oscillation as by it. A coil of large self-inductance and small resistance placed in an a.c. circuit makes ϕ approach $\pi/2$ and cuts down the current by increasing

the impedance without wasting power in the development of heat as a rheostat does in a d.c. circuit. Such a coil is called a *choke coil.*

The current given by (90–2) is that existing in the steady state. In addition we may have transients. These, however, are rapidly damped out if there is any appreciable resistance in the circuit.

Problem 90a. What value of the capacity will make the circuit of this article equivalent to that of article 87? Justify your answer physically.

Problem 90b. A circuit has an alternating e.m.f. in volts of 100 cos ωt, a resistance of 2 ohms, a self-inductance of 0.001 henry and a condenser of capacity 1000 microfarads, all connected in series. Find the effective current for the following values of ω: 1/sec, 10/sec, 100/sec, 1000/sec, 10,000/sec, 100,000/sec, 1,000,000/sec. Also find the frequency of resonance. Ans. 0.071 amp, 0.707 amp, 6.93 amp, 35.3 amp, 6.93 amp, 0.707 amp, 0.071 amp; 159 cycle.

Problem 90c. A 60-cycle a.c. circuit has resistance of 2 ohms and inductance of 10 millihenrys. What is the power factor? What capacity, placed in the circuit, will make the power factor unity? By how much does the insertion of this capacity increase the current? Ans. 0.47, 0.704 millifarad, 115%.

Problem 90d. Show that the mean power expended by the external e.m.f. in the circuit under consideration in this article is equal to the rate at which heat is developed in the resistance.

CHAPTER X

INTERACTION OF CURRENTS

91. Energy of an Isolated Circuit. — The equation (89–3) of a circuit containing self-inductance, capacity and resistance has the same form as that of a damped simple harmonic oscillator. Consider, for instance, a mass m (Fig. 197) suspended from O by means of a spring S. When displaced vertically a distance x from its equilibrium position it experiences a force of restitution $-kx$ due to the spring, a damping force $-b\frac{dx}{dt}$ on account of air resistance, and in addition it may be subject to an external force F. So its equation of motion is

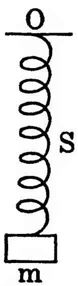


FIG. 197

$$m\frac{d^2x}{dt^2} + b\frac{dx}{dt} + kx = F.$$

Comparing this equation with (89–3), we see that the self-inductance of the circuit corresponds to the mass of the oscillator, the resistance to the damping constant b and the reciprocal of the capacity to the stiffness k of the spring. We can carry the analogy as far as we choose, likening, for instance, the quantity Li to the momentum of the oscillator. Thus, once the current is established, its momentum tends to keep it flowing even in the absence of an impressed e.m.f. This momentum, however, is dissipated in part in overcoming resistance and the remainder is used up in charging the condenser just as that of the oscillator is lost in extending or compressing the spring. After all the momentum has been absorbed and the current has ceased to flow, the condenser starts to discharge, giving rise to a current in the opposite sense, exactly as in the mechanical analog the

oscillating mass comes to rest after extending the spring and then starts to move in the opposite direction as the spring contracts.

The rate at which work W is done by the external e.m.f. in the circuit under consideration is obtained by multiplying (89–3) by $i = \frac{dq}{dt}$, giving

$$\frac{dW}{dt} = \mathcal{E}\frac{dq}{dt} = Li\frac{di}{dt} + Ri^2 + \frac{1}{C}q\frac{dq}{dt}.$$

Integrating from time t_1 to time t_2 the total work done in the time $t_2 - t_1$ is

$$W = \tfrac{1}{2}Li_2^2 - \tfrac{1}{2}Li_1^2 + \int_{t_1}^{t_2} Ri^2dt + \frac{q_2^2}{2C} - \frac{q_1^2}{2C}, \qquad (91\text{–}1)$$

where i_1 is the current in the circuit at the time t_1 and i_2 that at the time t_2, q_1 and q_2 being the charges on the condenser at the same times. As the integrand in the term involving R is always positive, this term becomes greater and greater the longer the interval of time. It represents an irreversible dissipation of energy, which is transformed into heat in the resistance. On the other hand, the remaining terms specify an increase of energy of the circuit which is reconverted into mechanical work when the current and charge resume their initial values. The quantity

$$\frac{1}{2}\frac{q^2}{C}$$

has already (art. 21) been shown to be the potential energy of the charged condenser. The expression

$$\tfrac{1}{2}Li^2$$

represents energy possessed by the circuit on account of its self-inductance. As self-inductance is a consequence of the magnetic field of the current, we can interpret this as *magnetic energy*. By analogy with the oscillator we may describe it as *kinetic* rather than *potential* in character. In fact it represents the kinetic energy of the electrons constituting the current.

In the case of an alternating current the magnetic energy is

$$U_H = \tfrac{1}{2}Li^2 = \frac{L\mathcal{E}_0^2}{2Z^2}\sin^2(\omega t - \phi)$$
$$= \frac{L\mathcal{E}_0^2}{4Z^2}\{1 - \cos 2(\omega t - \phi)\}$$

from (90–2). The charge on the condenser is

$$q = \int i dt = -\frac{\mathcal{E}_0}{Z\omega}\cos(\omega t - \phi),$$

the constant of integration vanishing since q is zero when i is maximum. So the electric energy is

$$U_E = \frac{1}{2}\frac{q^2}{C} = \frac{\mathcal{E}_0^2}{2CZ^2\omega^2}\cos^2(\omega t - \phi)$$
$$= \frac{\mathcal{E}_0^2}{4CZ^2\omega^2}\{1 + \cos 2(\omega t - \phi)\},$$

from which we see that the one energy is maximum when the other is minimum and *vice versa*. The energy of the circuit, therefore, passes from the kinetic to the potential form and back again with a frequency double that of the impressed e.m.f., just as in the analogous case of the mechanical oscillator. If we calculate the work done during a complete period we see from (91–1) that it is entirely accounted for by the heat developed in the resistance.

Problem 91a. At what impressed frequency is the mean kinetic energy equal to the mean potential energy in the circuit treated in the latter part of this article? Ans. Resonance.

Problem 91b. A circuit has inductance and capacity but no appreciable resistance. Show that when the circuit is oscillating freely (no external e.m.f.) the mean kinetic and potential energies are equal.

92. Energy of a System of Current Circuits. — In the last article we confined our attention to a single circuit. Now we shall consider the interactions of a number of circuits near one another and fixed relative to the observer. Although the

method to be developed is applicable to any number of circuits we shall confine the analysis to the case of three circuits designated by the numbers (1), (2) and (3). As we are interested only in the magnetic energy of the circuits we shall suppose that they contain no condensers.

The total flux of induction through (1) is made up of three parts, the flux due to the current i_1 in circuit (1) itself, the flux due to the current i_2 in circuit (2), and the flux due to the current i_3 in circuit (3). The first of these, as we have already seen, is $L_1 i_1$, where L_1 is the self-inductance of (1). As the magnetic field due to the current in (2) is proportional to i_2, the flux through (1) occasioned by it may be written $M_{12} i_2$. The coefficient M_{12} is a function of the coordinates specifying the relative positions of the two circuits, but is independent of i_2 provided the permeability of the medium surrounding the circuits is constant. It is smaller in absolute magnitude the farther apart the circuits are placed. We call M_{12} the *mutual inductance* of circuit (1) with respect to circuit (2). It represents the flux through (1) due to a unit current in (2). It is positive or negative according as the flux passes through the two circuits in the same or in opposite senses. Evidently mutual inductance is measured in the same units as self-inductance.

Finally we have the flux $M_{13} i_3$ through circuit (1) due to the current i_3 in (3), M_{13} being the mutual inductance of (1) with respect to (3). So the entire flux through circuit (1) is

$$N_1 = L_1 i_1 + M_{12} i_2 + M_{13} i_3, \qquad (92\text{–}1)$$

and if the currents change, the circuits being held fixed in position, an electromotive force

$$\mathcal{E}_1' = -\frac{dN_1}{dt} = -L_1 \frac{di_1}{dt} - M_{12}\frac{di_2}{dt} - M_{13}\frac{di_3}{dt} \qquad (92\text{–}2)$$

is induced in circuit (1), provided μ is constant. From this equation we can define M_{12} as the e.m.f. induced in (1) by a unit time rate of decrease of current in (2), and so on. Expressions

similar to (92–2) give the induced electromotive forces in circuits (2) and (3).

The magnetic energy of the system of circuits is equal to the work done against the induced e.m.f.'s as in the case of the single circuit discussed in the last article. Therefore we do not need to write down the complete equations of the circuits in order to calculate the magnetic energy of the system, but can obtain the desired result by computing the work done against the induced electromotive forces $\mathcal{E}_1'$, $\mathcal{E}_2'$, $\mathcal{E}_3'$ in the three circuits by the currents i_1, i_2, i_3 flowing in them. In a time dt this work is

$$dU = -\mathcal{E}_1' i_1 dt - \mathcal{E}_2' i_2 dt - \mathcal{E}_3' i_3 dt. \tag{92–3}$$

By starting with zero current in each circuit and gradually building the current up to its final value we can calculate the magnetic energy of the system of circuits. Denote by I_1, I_2, I_3 the final values of the currents in the three circuits and suppose that the currents increase in such a way that at any instant each current is the same fraction α of its final value. Then $i_1 = \alpha I_1$, $i_2 = \alpha I_2$, $i_3 = \alpha I_3$, where α increases from 0 to 1 during the process of building up the currents. Putting (92–2) and similar expressions for the electromotive forces in the other circuits into (92–3), we find for the total work done against the induced e.m.f.'s in establishing the currents

$$\begin{aligned} U = \{ & L_1 I_1^2 + M_{12} I_1 I_2 + M_{13} I_1 I_3 \\ & + M_{21} I_2 I_1 + L_2 I_2^2 + M_{23} I_2 I_3 \\ & + M_{31} I_3 I_1 + M_{32} I_3 I_2 + L_3 I_3^2 \} \int_{\alpha=0}^{\alpha=1} \alpha \frac{d\alpha}{dt} dt, \end{aligned}$$

giving

$$\left. \begin{aligned} U = \tfrac{1}{2}\{ & L_1 I_1^2 + M_{12} I_1 I_2 + M_{13} I_1 I_3 \\ & + M_{21} I_2 I_1 + L_2 I_2^2 + M_{23} I_2 I_3 \\ & + M_{31} I_3 I_1 + M_{32} I_3 I_2 + L_3 I_3^2 \}. \end{aligned} \right\} \tag{92–4}$$

This expression does not represent *all* the work which must be performed to establish the currents, but only the work done against the induced electromotive forces. In addition energy

has to be supplied to overcome the resistances R_1, R_2, R_3 of the circuits at the rate

$$R_1 i_1^2 + R_2 i_2^2 + R_3 i_3^2,$$

and this expenditure of energy continues after the currents have attained their final steady values. Nevertheless, the energy used in overcoming resistance is all dissipated in heat and does not add to the magnetic energy of the system of circuits, which is given completely by (92–4).

If, now, we vary the current I_1 keeping the other currents constant, the change in the magnetic energy of the circuits is

$$\frac{\partial U}{\partial I_1} dI_1 = \{L_1 I_1 + \tfrac{1}{2}(M_{12} + M_{21})I_2 + \tfrac{1}{2}(M_{13} + M_{31})I_3\} dI_1. \tag{92–5}$$

But when I_1 alone is changed,

$$\mathcal{E}_1' = - L_1 \frac{dI_1}{dt}, \qquad \mathcal{E}_2' = - M_{21} \frac{dI_1}{dt}, \qquad \mathcal{E}_3' = - M_{31} \frac{dI_1}{dt},$$

and the work done against the induced electromotive forces is

$$\begin{aligned} \frac{\partial U}{\partial I_1} dI_1 &= \left\{ L_1 I_1 \frac{dI_1}{dt} + M_{21} I_2 \frac{dI_1}{dt} + M_{31} I_3 \frac{dI_1}{dt} \right\} dt \\ &= \{L_1 I_1 + M_{21} I_2 + M_{31} I_3\}\, dI_1. \end{aligned} \tag{92–6}$$

Equations (92–5) and (92–6) are true whatever values we choose to give I_1, I_2 and I_3. If, for instance, we make I_1 and I_3 zero, then we find by equating (92–5) and (92–6) that $M_{12} = M_{21}$. Similarly we find that the other mutual inductances are equal in pairs. In all we have

$$M_{12} = M_{21}, \qquad M_{23} = M_{32}, \qquad M_{31} = M_{13}. \tag{92–7}$$

Hence we conclude that the flux through circuit (1) due to unit current in (2) is equal to the flux through (2) due to unit current in (1) and so on. Instead of six independent mutual inductances we have but three.

An isolated circuit in which a current i is flowing, then, has the magnetic energy

$$U = \tfrac{1}{2} L i^2, \tag{92–8}$$

and two neighboring circuits the energy

$$U = \tfrac{1}{2}L_1 i_1^2 + M i_1 i_2 + \tfrac{1}{2}L_2 i_2^2, \qquad (92\text{–}9)$$

the middle term representing the mutual energy of the pair. If we separate the circuits, keeping the currents constant, the mutual energy becomes smaller and smaller in absolute value, approaching zero for infinite separation.

In the case of three circuits we can write (92–4) in the simpler form

$$U = \tfrac{1}{2}(L_1 i_1^2 + L_2 i_2^2 + L_3 i_3^2) + M_{12} i_1 i_2 + M_{23} i_2 i_3 + M_{31} i_3 i_1, \qquad (92\text{–}10)$$

by virtue of the relations (92–7).

Problem 92a. Find the energy of the circuit of problem 87*a* if $l = 100$ cm, $a = 1$ cm, $b = 10$ cm, $i = 5$ amp, $\mu = 100$. Ans. $0.576(10)^4$ erg.

93. Forces and Torques on Rigid Circuits. — For simplicity we shall confine ourselves to a pair of circuits and calculate the force or torque on one due to the other. Each circuit will be supposed to be rigid, the only motion considered being that of the one relative to the other. If one circuit is fixed relative to the observer and the other suffers a small displacement $d\xi$, the work done by the magnetic force exerted by the first circuit on the second is $F d\xi$, where F is the component of the force in the direction of the displacement.

As each circuit is rigid the relative motion of the two circuits changes the mutual inductance M but not the self-inductances L_1 and L_2 of the two circuits. The change in mutual inductance means that the flux through each circuit changes, resulting in an induced e.m.f. and a change in current. The flux through the first circuit is

$$N_1 = L_1 i_1 + M i_2,$$

and the e.m.f. induced when both the currents and the mutual

inductance change as a result of the displacement is

$$\mathcal{E}_1' = -L_1\frac{di_1}{dt} - M\frac{di_2}{dt} - i_2\frac{dM}{dt}.$$

If, then, $\mathcal{E}_1$ is the external e.m.f. and R_1 the resistance of circuit (1), the equation of the circuit is

$$\mathcal{E}_1 + \mathcal{E}_1' = R_1 i_1,$$

or

$$L_1\frac{di_1}{dt} + M\frac{di_2}{dt} + i_2\frac{dM}{dt} + R_1 i_1 = \mathcal{E}_1. \qquad (93\text{–}1)$$

Writing down a similar expression for circuit (2), the rate at which work is done on the two circuits by the impressed electromotive forces $\mathcal{E}_1$ and $\mathcal{E}_2$ is

$$\begin{aligned}\frac{dW}{dt} &= \mathcal{E}_1 i_1 + \mathcal{E}_2 i_2 \\ &= L_1 i_1\frac{di_1}{dt} + L_2 i_2\frac{di_2}{dt} + M i_1\frac{di_2}{dt} + M i_2\frac{di_1}{dt} \\ &\qquad + 2i_1 i_2\frac{dM}{dt} + R_1 i_1^2 + R_2 i_2^2. \qquad (93\text{–}2)\end{aligned}$$

The rate of increase of magnetic energy obtained by differentiating (92–9) with respect to the time is

$$\frac{dU}{dt} = L_1 i_1\frac{di_1}{dt} + L_2 i_2\frac{di_2}{dt} + M i_1\frac{di_2}{dt} + M i_2\frac{di_1}{dt} + i_1 i_2\frac{dM}{dt},$$

where we have kept L_1 and L_2 constant since the circuits are rigid. Now the work performed per unit time by the external sources of energy is equal to the sum of the rate $F\frac{d\xi}{dt}$ at which mechanical work is done by the magnetic forces between the two circuits, the rate $\frac{dU}{dt}$ at which the magnetic energy of the pair of circuits increases, and the rate $R_1 i_1^2 + R_2 i_2^2$ at which energy

is dissipated in heat. Consequently

$$\frac{dW}{dt} = F\frac{d\xi}{dt} + L_1 i_1 \frac{di_1}{dt} + L_2 i_2 \frac{di_2}{dt} + M i_1 \frac{di_2}{dt} + M i_2 \frac{di_1}{dt} + i_1 i_2 \frac{dM}{dt} + R_1 i_1^2 + R_2 i_2^2. \quad (93\text{-}3)$$

Equating (93-2) and (93-3),

$$F\frac{d\xi}{dt} = i_1 i_2 \frac{dM}{dt}, \quad (93\text{-}4)$$

from which

$$F = i_1 i_2 \frac{\partial M}{\partial \xi}. \quad (93\text{-}5)$$

If we are concerned with a torque instead of a force, the rate at which mechanical work is done when one of the circuits rotates through an angle $d\theta$ is $L\frac{d\theta}{dt}$, where L is the torque about the axis of rotation. In this case we are led to the relation

$$L\frac{d\theta}{dt} = i_1 i_2 \frac{dM}{dt},$$

or

$$L = i_1 i_2 \frac{\partial M}{\partial \theta}, \quad (93\text{-}6)$$

for a pair of rigid circuits.

From these equations it is clear that the force or torque on the movable circuit is in such a direction as to increase M. Therefore a circuit tends to move in such a way as to increase the flux of induction through it, the flux being taken as positive when in the direction of advance of a right-handed screw rotated in the sense in which the current is flowing. The position of stable equilibrium of the circuit is that for which the flux is a maximum, subject to whatever mechanical constraints may be imposed.

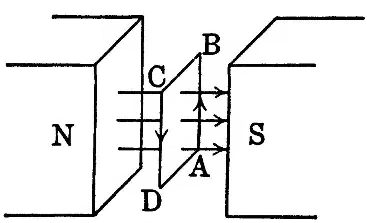


FIG. 198

If, for instance, a rectangular circuit $ABCD$ (Fig. 198) in which a current is flowing in the sense of the arrow-heads is free to rotate in the gap between the poles of a magnet, the position of stable equilibrium is that in which the positive face of the circuit is turned toward the south pole and the negative face toward the north pole, as in the illustration. Hence the positive face of the circuit acts like the north pole of a magnet and the negative face like a south pole.

As the flux through the movable circuit (1) due to the other (2) is

$$N = Mi_2,$$

we can write (93–5) and (93–6) in the forms

$$F = i_1 \frac{\partial N}{\partial \xi}, \tag{93–7}$$

$$L = i_1 \frac{\partial N}{\partial \theta}, \tag{93–8}$$

where the change in flux represented by the derivative is that due to the motion of the circuit alone and does not include any change in flux produced by an alteration in the intensity of the field. In other words, the change in flux involved is that which takes place when the current i_2 in the fixed circuit is kept constant.

Since the force or torque on the movable circuit may be considered as due to the magnetic field in its neighborhood, it makes no difference whether the flux N is produced by a second circuit as we have supposed or by a magnet. In fact it may be due to a number of neighboring circuits and magnets acting together. Therefore formulas (93–7) and (93–8) are quite general in their applicability.

If the permeability of the medium in which the circuits are immersed is unity, $B = H$ and N in these equations is the flux of magnetic intensity. Comparing with (41–5) and (41–6), we see that the force or torque on a current circuit in empty space is the same as that on a magnetic shell whose periphery coincides with the circuit and whose strength is equal to the current. In article 71 it was shown that such a magnetic shell produces the

same field as the current at all outside points. Therefore we have established the equivalence of a magnetic shell and a current circuit both as regards the field produced and the force or torque experienced under the action of an external field.

On the other hand, if a surrounding medium of permeability different from unity is present, N in (93–7) and (93–8) represents the flux of B whereas N in (41–5) and (41–6) stands for the flux of H. Nevertheless the normal component of H in the disk-shaped cavity in the medium occupied by the shell is equal to the normal component of B in the medium just outside the shell, and therefore to the component of B normal to the surface bounded by the circuit in the absence of the equivalent shell. So in this case too we have an exact equivalence both in the field produced and in the force or torque experienced.

Finally let us return to the problem of the interaction of two current circuits and consider the case where the currents are kept constant during the displacement $d\xi$ of circuit (1). In order to keep the current in (1) constant it is necessary to apply an external electromotive force $\mathcal{E}_1$ (in addition to that necessary to overcome resistance) equal and opposite to the e.m.f. $\mathcal{E}_1'$ induced by the changing flux. As i_1 and i_2 do not vary,

$$\mathcal{E}_1' = -i_2\frac{dM}{dt},$$

and

$$\mathcal{E}_1 = -\mathcal{E}_1' = i_2\frac{dM}{dt}.$$

The rate at which work is done by this external e.m.f. is

$$\mathcal{E}_1 i_1 = i_1 i_2\frac{dM}{dt}.$$

As work is done by the external e.m.f. necessary to keep the current constant in circuit (2) at an equal rate, the total energy which must be supplied from outside per unit time is

$$2i_1 i_2\frac{dM}{dt}. \qquad (93\text{–}9)$$

This, however, is double the rate (93–4) at which work is done by the mechanical force acting on the movable circuit. Therefore, if the currents in the two circuits are kept constant, the magnetic energy of the field increases at the same rate as that at which mechanical work is performed. This is known as *Kelvin's law.* The phenomenon is quite analogous to that discussed in article 22 for the case of charged conductors, where it was shown that if the potentials of a set of conductors are kept constant during a displacement of one of them the electrostatic energy of the system increases at the same rate as that at which mechanical work is done, the sources of potential furnishing a double amount of energy.

Problem 93a. The center of a plane circular circuit of radius a consisting of one turn lies on the X axis at a distance x from the origin, the positive normal to the circuit being in the X direction. Find the force exerted on it by a radial field diverging from the origin of intensity $H = m/r^2$. Ans. $-\dfrac{2\pi mia^2}{(a^2+x^2)^{3/2}}$.

Problem 93b. Find the torque on the circuit of Fig. 198 for any angle θ between the normal to the circuit and the lines of force of the field. Assume that the coil has n turns of area A and that $\mu = 1$. Ans. $L = -nAHi \sin\theta$.

Problem 93c. The center O of a circular current of radius a lies at a distance R from the center C of a thin bar magnet of pole strength m and length l, the line OC being parallel to the axis of the magnet and perpendicular to the plane of the circuit. Find the force between them, R being greater than $l/2$.

$$\text{Ans.}\quad -2\pi ma^2 i\left[\frac{1}{\left\{\left(R-\frac{l}{2}\right)^2+a^2\right\}^{3/2}}-\frac{1}{\left\{\left(R+\frac{l}{2}\right)^2+a^2\right\}^{3/2}}\right].$$

94. Calculation of Inductances. — Perhaps the most important example of the action of one circuit on another is that taking place between two solenoids one of which is wound outside of the other as in Fig. 199. Suppose that the inner solenoid (*primary*) is very long whereas the outer (*secondary*) consists of a number of turns situated near the center of the primary. If i

is the current in the primary and n_1 the number of turns per unit length, the field far from the ends is

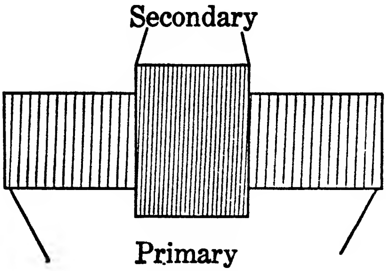


FIG. 199

$$H = 4\pi n_1 i$$

from (70–5), and if μ is the permeability of the core on which the primary is wound,

$$B = \mu H = 4\pi\mu n_1 i.$$

Consequently, if A is the cross-section of the primary, the flux through each turn of the secondary is

$$N_1 = BA = 4\pi\mu n_1 Ai,$$

and if the secondary has m turns the total flux linked is

$$N = mN_1 = 4\pi\mu m n_1 Ai.$$

Hence the mutual inductance of the two coils is

$$M = 4\pi\mu m n_1 A. \tag{94–1}$$

This gives the flux through either coil due to unit current in the other. Observe that M is independent of the cross-section of the secondary, for the field due to i is negligibly small outside the primary, and the flux through the secondary due to the current in the primary is not changed appreciably by making the cross-section of the secondary greater than that of the primary. While the mutual inductance depends upon the number of turns per unit length of the primary, it depends upon the total number of turns of the secondary.

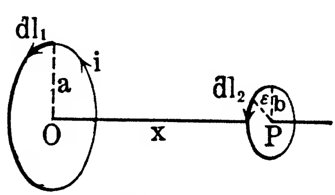


FIG. 200

A pair of circuits for which the approximate value of the mutual inductance is easily calculated consists of two parallel coaxial circles (Fig. 200) a distance x apart, one of which has a radius b small compared to the

radius a of the other. If a current i flows in the larger circuit, the field at the center P of the smaller is

$$H = \frac{2\pi i a^2}{(a^2 + x^2)^{3/2}}$$

from (70–3). As a fair approximation we can take this to represent the field over the entire cross-section of the smaller circle. Therefore the flux of induction through it is

$$N = \pi b^2 \mu H = \frac{2\pi^2 \mu a^2 b^2}{(a^2 + x^2)^{3/2}} i,$$

and the mutual inductance of the circuits is

$$M = \frac{2\pi^2 \mu a^2 b^2}{(a^2 + x^2)^{3/2}}. \qquad (94\text{–}2)$$

If the coils contain more than one turn this expression must be multiplied by the product of the number of turns in the two circuits.

In deducing (94–2) the currents have been supposed to flow in the same sense. This makes the mutual inductance, and therefore the flux through either circuit due to the other, positive in sign. Therefore, as the circuits tend to move in such a direction as to increase the flux, the two currents attract each other. In general, *parallel currents flowing in the same direction attract, in opposite directions repel.*

In problem 87b the student has calculated the self-inductance of a ring solenoid. We can deduce an approximate expression for the self-inductance of a straight solenoid of radius a and length l (Fig. 201) having n_1 turns per unit length if we suppose that the field is uniform across the cross-section of the solenoid and has the value given by (70–4) for a point on the axis. Then

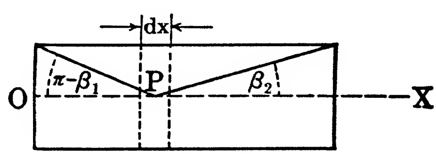


FIG. 201

at a point P at a distance x from the left end of the coil,

$$H = 2\pi n_1 i \{\cos \beta_2 + \cos (\pi - \beta_1)\}$$
$$= 2\pi n_1 i \left\{ \frac{x}{\sqrt{x^2 + a^2}} + \frac{(l - x)}{\sqrt{(l - x)^2 + a^2}} \right\},$$

and the flux through the $n_1 dx$ turns in a length dx is

$$dN = \pi a^2 \mu H n_1 dx$$
$$= 2\pi^2 \mu n_1^2 a^2 i \left\{ \frac{x dx}{\sqrt{x^2 + a^2}} + \frac{(l - x) dx}{\sqrt{(l - x)^2 + a^2}} \right\}.$$

Integrating from 0 to l,

$$N = 2\pi^2 \mu n_1^2 a^2 i \left\{ \left| \sqrt{x^2 + a^2} \right|_0^l - \left| \sqrt{(l - x)^2 + a^2} \right|_0^l \right\}$$
$$= 4\pi^2 \mu n_1^2 a^2 i \{\sqrt{l^2 + a^2} - a\}.$$

This gives for the self-inductance of the solenoid

$$L = 4\pi^2 \mu n_1^2 a^2 \{\sqrt{l^2 + a^2} - a\}, \tag{94-3}$$

an expression which is correct to within 2% if l is greater than $10a$. For solenoids which are shorter in proportion to their radii the formula

$$L = 4\pi^2 \mu n_1^2 a^2 l K \tag{94-4}$$

may be used, where K is a function of a/l which decreases from 1 to 0 as the argument increases from 0 to ∞ in accord with the curve of Fig. 202. A table of values of K may be found in *Circular 74 of the Bureau of Standards*, p. 283.

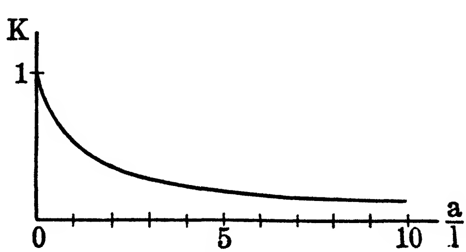


Fig. 202

A general expression which is fundamental in the calculation of inductances has been given by Neumann. To deduce this formula we shall first compute the force exerted by one circuit on another. Let dl_1 and dl_2

(Fig. 203) be elements of two circuits carrying currents i_1 and i_2 respectively. Denote by θ_1 and θ_2 the angles which the two elements make with the line $r = OP$ joining them. The magnetic field at P due to the first is

$$dH = \frac{i_1 dl_1 \sin \theta_1}{r^2}$$

directed out from the paper.

We shall calculate the force which this field exerts on the second current element. First resolve dl_2 into components $dl_2 \cos \theta_2$ along r and $dl_2 \sin \theta_2$ at right angles to r. If ϕ is the angle which the latter makes with the XOP plane we can resolve it into a component $dl_2 \sin \theta_2 \sin \phi$ parallel to dH and a component $dl_2 \sin \theta_2 \cos \phi$ perpendicular to dH in the XOP plane. No force is exerted by the field on the first of these since it is parallel to the lines of force.

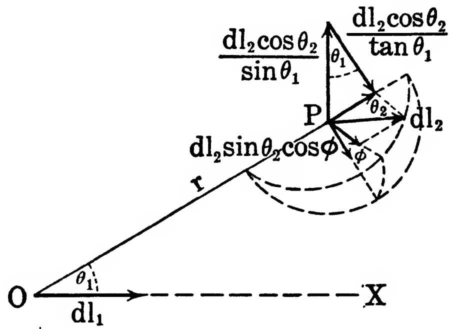


FIG. 203

Next resolve $dl_2 \cos \theta_2$ into a component $dl_2 \cos \theta_2/\sin \theta_1$ perpendicular to dl_1 and a component $dl_2 \cos \theta_2/\tan \theta_1$ perpendicular to OP. Then, neglecting the component parallel to dH since it is subject to no force, we have

$$\frac{dl_2 \cos \theta_2}{\sin \theta_1}$$

perpendicular to dl_1, and

$$dl_2 \left(\frac{\cos \theta_2}{\tan \theta_1} + \sin \theta_2 \cos \phi \right)$$

perpendicular to OP. The forces on these two current elements are

$$d^2F_x = i_1 i_2 \frac{dl_1 \sin \theta_1}{r^2} \frac{dl_2 \cos \theta_2}{\sin \theta_1} = i_1 i_2 \frac{dl_1 dl_2 \cos \theta_2}{r^2}$$

$$= i_1 i_2 dl_1 \frac{dr}{r^2} \tag{94-5}$$

in the X direction, since $dr = dl_2 \cos \theta_2$, and

$$d^2F_r = -\, i_1 i_2 \frac{dl_1 \sin \theta_1}{r^2}\, dl_2 \left(\frac{\cos \theta_2}{\tan \theta_1} + \sin \theta_2 \cos \phi \right)$$

$$= -\, i_1 i_2 \frac{dl_1 dl_2}{r^2} (\cos \theta_1 \cos \theta_2 + \sin \theta_1 \sin \theta_2 \cos \phi) \tag{94-6}$$

in the direction of the radius vector OP.

Now if we integrate (94–5) around the circuit (2) we get

$$dF_x = i_1 i_2 dl_1 \oint \frac{dr}{r^2} = i_1 i_2 dl_1 \left| -\frac{1}{r} \right| = 0,$$

as the limits coincide. Therefore we have left only the force (94–6) along the radius vector. The trigonometrical expression in the parentheses is just the cosine of the angle ϵ between dl_1 and dl_2. Therefore, introducing vector notation,

$$d^2F_r = -\, i_1 i_2 \frac{dl_1 dl_2 \cos \epsilon}{r^2} = -\, i_1 i_2 \frac{d\mathbf{l}_1 \cdot d\mathbf{l}_2}{r^2}$$

$$= i_1 i_2 \frac{\partial}{\partial r} \left(\frac{d\mathbf{l}_1 \cdot d\mathbf{l}_2}{r} \right).$$

Consequently the rate at which work is done if we displace one circuit relative to the other is

$$d^2\mathcal{P} = d^2F_r \frac{dr}{dt} = i_1 i_2 \frac{d}{dt} \left(\frac{d\mathbf{l}_1 \cdot d\mathbf{l}_2}{r} \right),$$

and integrating first around one circuit and then around the other

$$\mathcal{P} = i_1 i_2 \frac{d}{dt} \oint \oint \frac{d\mathbf{l}_1 \cdot d\mathbf{l}_2}{r}.$$

Comparing with (93–4) we obtain *Neumann's formula*,

$$M = \oint \oint \frac{d\mathbf{l}_1 \cdot d\mathbf{l}_2}{r}. \tag{94-7}$$

The symmetry of this expression shows at once that two circuits have but a single mutual inductance, i.e. $M_{12} = M_{21}$. This formula enables us to calculate not only the mutual induc-

tance of two circuits but also the self-inductance of a single circuit. For we can consider a circuit to be made up of a bundle of current filaments, the self-inductance being the mutual inductance of the filaments with respect to one another.

By the use of Neumann's formula we may obtain a closer approximation than (94–2) for the mutual inductance of two parallel coaxial circles (Fig. 200) of radii a and b without limiting ourselves to the case where b is small compared to a. From the figure we have

$$d\mathbf{l}_1 \cdot d\mathbf{l}_2 = dl_1 dl_2 \cos \epsilon = dl_1 b \cos \epsilon d\epsilon = b dl_1 \left(2 \cos^2 \frac{\epsilon}{2} - 1 \right) d\epsilon,$$

and the distance r between dl_1 and dl_2 is given by

$$\begin{aligned} r^2 &= x^2 + (a - b \cos \epsilon)^2 + b^2 \sin^2 \epsilon \\ &= x^2 + (a + b)^2 - 4ab \cos^2 \frac{\epsilon}{2} \cdot \end{aligned}$$

Putting

$$h^2 \equiv x^2 + (a + b)^2,$$

we have

$$r^2 = h^2 \left(1 - \frac{4ab}{h^2} \cos^2 \frac{\epsilon}{2} \right),$$

and expanding by the binomial theorem,

$$\frac{1}{r} = \frac{1}{h} \left(1 + 2 \frac{ab}{h^2} \cos^2 \frac{\epsilon}{2} + 6 \frac{a^2 b^2}{h^4} \cos^4 \frac{\epsilon}{2} + 20 \frac{a^3 b^3}{h^6} \cos^6 \frac{\epsilon}{2} + \cdots \right) \cdot$$

Therefore

$$\begin{aligned} M = \frac{b}{h} \oint dl_1 \oint \left(2 \cos^2 \frac{\epsilon}{2} - 1 \right) d\epsilon \bigg(1 &+ 2 \frac{ab}{h^2} \cos^2 \frac{\epsilon}{2} \\ &+ 6 \frac{a^2 b^2}{h^4} \cos^4 \frac{\epsilon}{2} + 20 \frac{a^3 b^3}{h^6} \cos^6 \frac{\epsilon}{2} + \cdots \bigg) \cdot \end{aligned}$$

Integrating term by term,

$$\begin{aligned} M &= \frac{\pi a b^2}{h^3} \oint dl_1 \left(1 + 3 \frac{ab}{h^2} + \frac{75}{8} \frac{a^2 b^2}{h^4} + \cdots \right) \\ &= \frac{2\pi^2 a^2 b^2}{h^3} \left(1 + 3 \frac{ab}{h^2} + 9.375 \frac{a^2 b^2}{h^4} + \cdots \right) \cdot \qquad (94\text{–}8) \end{aligned}$$

If b is neglected as compared to a, the first term of (94–8) is identical with (94–2) with the omission of the factor μ which has been taken to be unity in the present calculation.

Problem 94a. An air solenoid of 2 cm radius and 50 cm length has 400 turns. A secondary of 200 turns, radius 2.5 cm and length 25 cm is wound around it so that the centers of the two windings coincide. Find the self-inductance of each coil and the mutual inductance of the pair. Ans. 0.48 millihenry, 0.37 millihenry, 0.25 millihenry.

Problem 94b. A ring of mean radius a and cross-sectional radius b is surrounded by two windings; a primary of n_p turns and a secondary of n_s turns. Find the mutual inductance of the two circuits. (See problem 87*b*.) Ans. $M = 4\pi\mu n_p n_s(a - \sqrt{a^2 - b^2})$.

Problem 94c. Two parallel coaxial circles of the same radius a are a distance $7.75a$ apart. Compute their mutual inductance. Ans. $0.0405a$.

Problem 94d. Find the force between two circular coaxial currents i_1 and i_2 of radii a and b a distance x apart.

$$\text{Ans.}\quad -\frac{6\pi^2a^2b^2i_1i_2}{h^5}x\left\{1 + 5\frac{ab}{h^2} + 21.875\frac{a^2b^2}{h^4} + \cdots\right\}.$$

95. Examples of Electromagnetic Reactions. — Here we shall discuss some examples of the theory developed in the preceding articles. First we shall deduce Lenz's law.

Lenz's Law. — Consider a turn of wire placed in a magnetic field. If we move it in such a way as to increase the flux through it as illustrated in Fig. 204*a*, a current is induced in such a sense

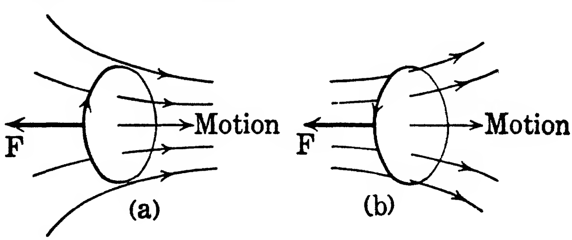


Fig. 204

as to give rise to an oppositely directed flux. Therefore the left-hand face of the circuit is positive relative to the sense in which the current is flowing and the force exerted on the circuit by the

magnetic field is in the direction to produce a displacement which will increase the flux from right to left or decrease the flux from left to right. That is, the magnetic force is in the opposite sense to the motion. The same is true if the circuit is moved so as to decrease the flux as in Fig. 204*b*, and the reader can easily show that similar conclusions are reached regarding the torque in the case of rotation. *In all cases the induced current is in such a direction as to oppose the motion which generates it.* This is known as *Lenz's law.*

A striking application of Lenz's law is afforded by the copper pendulum (Fig. 205). This consists of a copper plate C suspended from O so that it can swing back and forth between the poles N, S of a large electromagnet. The pendulum is set swinging while the magnet is unexcited. If, now, the switch controlling the current in the magnet coils is closed, induced currents — known as *eddy currents* — are set up in the moving copper plate in such directions as to oppose the motion of the pendulum. Consequently the latter, which may be moving with considerable speed at the instant the magnet is excited, is brought almost to a stand-still, and slowly sinks to its equilibrium position.

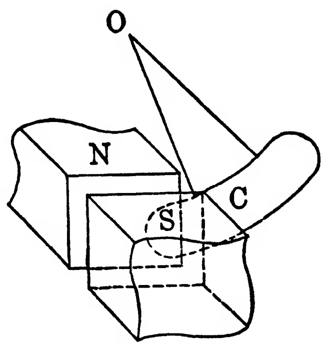


FIG. 205

The energy of the pendulum is converted into heat through the agency of the eddy currents set up by its motion through the magnetic field. These eddy currents have the further effect of giving rise to a flux opposed to that of the magnet and thereby diminishing the flux which passes through the surface of the copper plate. This shielding effect is greater the less the resistance of the plate, a perfect conductor constituting a complete shield against variations in the magnetic field.

Electromagnetic Shielding. — To illustrate more fully the phenomenon of *electromagnetic shielding*, consider a solenoid L

(Fig. 206) through which an alternating current $i_0 \sin \omega t$ is flowing and a neighboring loop of wire A near to which is placed a parallel metal plate C. If we suppose the latter to be a perfect conductor, the e.m.f. around every closed curve lying entirely inside it must vanish. Therefore the change of flux through such a curve due to the varying current in L must be compensated by an equal and opposite change of flux produced by the eddy currents in the plate. As the resultant variation of flux is zero inside the plate, it is very small near to the plate. So if A is placed close to C — it makes little difference whether it is on the one side or the other — there is very little variation of flux through A and therefore only a very small induced e.m.f.

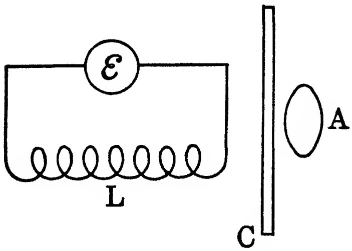


Fig. 206

If instead of a finite plate C a perfectly conducting shell is constructed so as to entirely surround L, the region outside is completely shielded from the varying flux produced by the alternating current in the solenoid. Actually the shielding is never complete since even the best conductors have some resistance.

On account of the dissipation of energy in heat through the production of eddy currents it is important to have no large pieces of metal in the neighborhood of electrical machines or instruments which make use of e.m.f.'s induced by varying flux. Both to avoid electromagnetic shielding and to reduce eddy currents to a minimum the iron cores used in transformers and in the armatures of a.c. generators and motors are laminated, that is, cut into disks or sheets insulated from one another by shellac or paint. The particular type of lamination employed in each case is that which gives the least opportunity for the generation of eddy currents. Thus to reduce eddy currents in the copper pendulum we might laminate the plate into narrow strips parallel to its longest edge.

Direct Current Motor. — In the simplest type of d.c. motor an armature consisting of a number of rectangular coils wound at different orientations around a cylindrical soft iron core rotates in the space between the poles of a large magnet. By means of a commutator the current is allowed to pass through only the one coil which is in a position to exert the greatest torque at the instant considered, that is, the one which has its face parallel to the field. We need, therefore, to consider only one rectangular coil *CDEF* (Fig. 207) rotating about an axis *PQ* between the poles *N*, *S* of the field magnet. If it contains n turns and has an area A the flux through it due to the external field when the normal to the plane of the coil makes an angle θ with the lines of force is

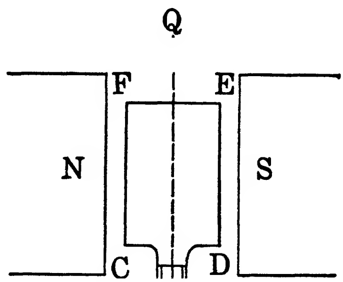


FIG. 207

$$N = nAB\cos\theta, \qquad (95\text{–}1)$$

and the torque is

$$\frac{dN}{d\theta} = -\,niAB\sin\theta$$

from (93–8), which is greatest when $\theta = \pi/2$, as stated above. The steady torque on the armature, then, is

$$L = -\,niAB \qquad (95\text{–}2)$$

in dyne cm if i and B are given in e.m.u. The minus sign indicates that the torque is in the opposite sense to that in which θ is measured.

On account of its motion through the field, however, an e.m.f.

$$\mathcal{E}_N = -\frac{dN}{dt} = \qquad \sin\theta$$

is induced in the coil. As the coil under consideration is connected to the external circuit through the commutator only

when $\theta = \pi/2$ we have a steady e.m.f. generated of amount

$$\mathcal{E}_N = nAB\omega. \tag{95-3}$$

As the coil rotates in the sense of the torque, $\omega = \dfrac{d\theta}{dt}$ is negative and the induced e.m.f. opposes the current, as might have been predicted from Lenz's law. Since $\mathcal{E}_N$ is negative, it is more convenient to consider the positive e.m.f. $\mathcal{E}_N' = -\mathcal{E}_N$ acting in the opposite sense to the current. This e.m.f. is called the *counter electromotive force* of the motor. If $\mathcal{E}$ is the impressed e.m.f. and R the resistance of the circuit, we have

$$\mathcal{E} - \mathcal{E}_N' = Ri \tag{95-4}$$

from Ohm's law. If R is small the current is large when the motor is starting up since ω and therefore $\mathcal{E}_N'$ are then small. To avoid an excessive current a starting rheostat is placed in the circuit. By this means an additional resistance is connected in the circuit until the motor has come up to speed. If the motor is carrying no load and the frictional resistance to the motion of the armature is negligible, the speed will increase until $\mathcal{E}_N'$ becomes equal to $\mathcal{E}$ and i vanishes.

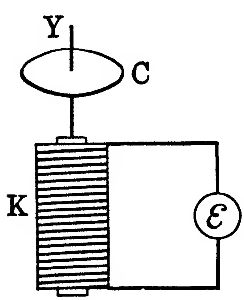


FIG. 208

Energy is developed by the impressed e.m.f. at the rate $\mathcal{E}i$, converted into mechanical work in the motor at the rate $\mathcal{E}_N'i$ and dissipated in heat at the rate Ri^2. To secure great efficiency the motor must be run at high speed so that i may be small and the resistance of the circuit must be as little as possible.

Electromagnetic Repulsion. — Consider a circular ring C (Fig. 208) of heavy copper wire placed above an electromagnet K which is connected with an alternating electromotive force $\mathcal{E}$. Denoting the current in K by

$$i_2 = i_0 \cos \omega t$$

and the mutual inductance of the two circuits by M, the flux through C due to the electromagnet is

$$N = Mi_2.$$

The induced e.m.f. in the ring, then, is

$$\mathcal{E}_1 = -\frac{dN}{dt} = M\omega i_0 \sin \omega t,$$

and according to (87–12) this gives rise to a current

$$i_1 = \frac{M\omega i_0}{Z_1} \sin (\omega t - \phi), \qquad \tan \phi \equiv \frac{L_1\omega}{R_1},$$

in C, where L_1 is the self-inductance, R_1 the resistance and Z_1 the impedance of the ring. Making use of (93–5) the upward force on C is

$$F = \frac{M\omega i_0^2}{Z_1} \sin (\omega t - \phi) \cos \omega t \frac{\partial M}{\partial y}. \tag{95–5}$$

As the distance y of the ring from the electromagnet increases, M decreases. Therefore $\frac{\partial M}{\partial y}$ is negative and we may write

$$F = -A \sin (\omega t - \phi) \cos \omega t,$$

where A is a positive quantity. Averaging over a period,

$$\begin{aligned} \sin (\omega t - \phi) \cos \omega t &= \sin \omega t \cos \omega t \cos \phi - \cos^2 \omega t \sin \phi \\ &= -\tfrac{1}{2} \sin \phi. \end{aligned}$$

Therefore the mean force is positive. If R_1 is small and ω large it may be of such a magnitude that the ring is thrown several feet up into the air when the alternating current is turned into the magnet coil. Decreasing R_1 increases the force both by making the impedance of the ring less and the lag ϕ greater.

Problem 95a. How should the armature core of the d.c. motor described in this article be laminated in order to minimize the production of eddy currents?

Problem 95b. Find the torque on the moving coil galvanometer of art. 74 by the method used here.

Problem 95c. Each coil in the armature of a d.c. motor consists of 100 turns of 200 cm² cross-section. The field is 4000 gauss and the external e.m.f. 110 volt. Find the counter e.m.f. at 1200 r.p.m. If the resistance of the circuit is 1 ohm what power is developed? Ans. 100.5 volt, 955 watt.

Problem 95d. The ring C of Fig. 208 has a radius of 12 cm, self-inductance of $(10)^{-7}$ henry and resistance $(10)^{-4}$ ohm. The electromagnet has a core of permeability 500 and cross-section 10 cm² wound with 10 turns per cm. The 60-cycle alternating current passing through it has an amplitude of 10 amp. The ring is placed 5 cm above the end of the magnet. Assuming that the magnet is so long that the effect of the distant pole can be neglected, that the intensity of magnetization is uniform and equal to that calculated for its mid-point, and that the pole adjacent to C can be treated as if concentrated on the axis, calculate the mean force on the ring. Ans. 5.8 lb wt.

96. Galvanometer Damping. — A single loop of heavy copper wire is generally attached to the moving part of a D'Arsonval galvanometer with its plane parallel to the coil. It is not connected electrically to the coil, but is provided for the purpose of damping the oscillations of the moving part through the agency of the currents induced by its rotation in the field of the fixed magnet. We shall discuss the motion of the suspended part of the instrument when the galvanometer is on open circuit so that there is no current in the coil.

If A is the area of the loop, H the field in which it swings, and α the angle which the plane of the loop makes with the lines of force, the flux through the loop is $N = AH \sin \alpha$. As the loop rotates an induced e.m.f.

$$\mathcal{E} = -\frac{dN}{dt} = -AH \cos \alpha \frac{d\alpha}{dt}$$

is set up in it. In calculating the induced current we can omit the term in the equation of the circuit which involves the self-inductance of the loop, for the time rate of change of current is so small for the slow oscillations we are concerned with here that this term is inappreciable as compared with the others. Hence the current is the quotient of $\mathcal{E}$ by the resistance R of the loop,

that is,

$$i = -\frac{AH}{R}\cos\alpha\frac{d\alpha}{dt}.$$

The torque due to the induced current is

$$L = i\frac{dN}{d\alpha} = -\frac{(AH)^2}{R}\cos^2\alpha\frac{d\alpha}{dt}, \qquad (96\text{–}1)$$

the negative sign indicating that the torque opposes the motion which produces it, in accord with Lenz's law.

In addition the moving coil is subject to a torque of restitution $-k\alpha$ due to the torsion of the suspension and a torque $-b\dfrac{d\alpha}{dt}$ due to air damping. So if I is the moment of inertia of the moving part, the equation of motion is

$$I\frac{d^2\alpha}{dt^2} = -k\alpha - b\frac{d\alpha}{dt} - \frac{(AH)^2}{R}\cos^2\alpha\frac{d\alpha}{dt}.$$

If the deflection is small or the field radial (art. 74) we can replace $\cos^2\alpha$ by unity. Then,

$$I\frac{d^2\alpha}{dt^2} + \left\{b + \frac{(AH)^2}{R}\right\}\frac{d\alpha}{dt} + k\alpha = 0. \qquad (96\text{–}2)$$

Since this equation is linear with constant coefficients the solution is of the form

$$\alpha = \alpha_0 e^{\gamma t},$$

where we find, by substituting in (96–2), that γ is given by

$$\gamma = -\frac{b + \dfrac{(AH)^2}{R}}{2I} \pm \frac{\sqrt{\left\{b + \dfrac{(AH)^2}{R}\right\}^2 - 4Ik}}{2I}.$$

As ordinarily designed the constants of the instrument are such that the expression under the radical is negative. Consequently the two values of γ are complex, and, as shown in article 89, the solution is

$$\alpha = \alpha_0 e^{-lt}\sin(\omega_0 t + \epsilon), \qquad (96\text{–}3)$$

where the damping constant l is

$$l \equiv \frac{1}{2I}\left\{b + \frac{(AH)^2}{R}\right\}. \tag{96–4}$$

The moving part of the galvanometer, then, oscillates about its position of equilibrium with rapidly decreasing amplitude. Since air resistance is small, by far the greater part of the damping is due to electromagnetic action in the loop. That this is so can easily be observed by removing the loop from the instrument, when it is found that the coil, upon being set into motion, takes much longer to come to rest. If the coil, instead of being on open circuit, is connected to a constant external e.m.f., the damping under consideration brings it rapidly to rest in the position for reading. In this case induced currents are set up in the coil as well as in the loop, thus increasing the electromagnetic damping.

If the loop is removed and the instrument short-circuited, electromagnetic damping is produced by the currents induced in the oscillating coil alone. If the coil has n turns of area A, (96–4) becomes

$$l \equiv \frac{1}{2I}\left\{b + \frac{(nAH)^2}{R}\right\}, \tag{96–5}$$

where R is the resistance of the coil and the external circuit. The coefficient $(nAH)^2$ of $1/R$ is just $(k/K_i)^2$ in terms of the symbols of (74–8). Although the resistance of the coil of a D'Arsonval galvanometer is generally about 100 ohms, the number of turns is so great that the electromagnetic damping due to the short-circuited coil alone is of the same order of magnitude as that due to the loop. By putting different external resistances in series with the coil after the loop has been removed and determining l experimentally, the straight line of Fig. 209 is

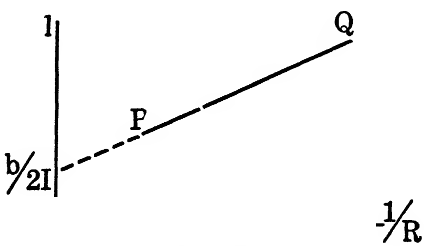


Fig. 209

obtained, confirming the formula (96–5) for l. Extending PQ to the axis of ordinates, the damping due to air resistance alone is obtained.

So far we have confined our attention to the moving coil galvanometer. Now we shall calculate the damping of the moving needle type. The change in flux through the fixed coil to be considered in this case is that due to the motion of the needle at its center. We can neglect the flux inside the needle since it does not change as the needle moves. First, then, we must compute the flux N_c through a circle AB (Fig. 210) of radius a due to a pole m lying near the axis of the circle at a distance h from its plane. The field at P is

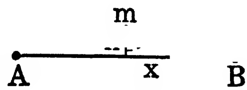


FIG. 210

$$m \qquad m$$

and, as we want the flux from below to above,

$$N_c = -\int (H \cos \theta) 2\pi x dx = -2\pi h^2 \int H \tan \theta \sec \theta d\theta,$$

since $x = h \tan \theta$. Consequently

$$N_c = -2\pi m \int \sin \theta d\theta = -2\pi m \left\{ 1 - \frac{h}{\sqrt{a^2 + h^2}} \right\}.$$

The south pole of the needle being an equal distance below , the total flux through the n turns in the coil is

$$N = -$$

and the induced e.m.f. is

$$\mathcal{E} = -\frac{dN}{dt} = - \qquad a^2 \qquad dh$$

$$= -\frac{4\pi nm\, dh}{a}$$

since h is negligible compared to a. Now if l is the length of the needle and α the angle which it makes with the plane of the coil,

$$h = \frac{l}{2}\sin\alpha, \qquad \frac{dh}{dt} = \frac{l}{2}\cos\alpha\frac{d\alpha}{dt},$$

and

$$\mathcal{E} = -\frac{2\pi nM}{a}\cos\alpha\frac{d\alpha}{dt},$$

where M is the magnetic moment ml of the needle. The current induced by this e.m.f. in the fixed coil is

$$i = \frac{\mathcal{E}}{R} = -\frac{2\pi nM}{aR}\cos\alpha\frac{d\alpha}{dt},$$

where R is the resistance of the circuit.

This gives rise to a field

$$H = \frac{2\pi ni}{a} = -\frac{4\pi^2n^2M}{a^2R}\cos\alpha\frac{d\alpha}{dt}$$

at the center of the coil, which exerts a damping torque

$$L = HM\cos\alpha = -\left(\frac{2\pi nM}{a}\right)^2\frac{1}{R}\cos^2\alpha\frac{d\alpha}{dt} \qquad (96\text{–}6)$$

on the needle.

As this equation takes the place of (96–1) in the case of the moving coil instrument, (96–5) must be replaced by

$$l \equiv \frac{1}{2I}\left\{b + \left(\frac{2\pi nM}{a}\right)^2\frac{1}{R}\right\}. \qquad (96\text{–}7)$$

The coefficient $(2\pi nM/a)^2$ of $1/R$ is the ratio $(k/K_i)^2$ of the symbols appearing in (74–7). Expressed in terms of the galvanometer constant, then, equation (96–7) is identical with (96–5).

Although the form of the expression for l is the same as in the case of the D'Arsonval galvanometer, the values of the quantities involved are such that electromagnetic damping is generally negligible in this type of instrument.

Problem 96a. The coil of a D'Arsonval galvanometer has n turns of 5 cm^2 area in a field of 500 gauss. The needle of a tangent galvanometer has a magnetic moment of 50 e.m.u., the fixed coil having n turns of 10 cm radius. Assuming the moment of inertia of the moving part and the resistance of the circuit to be the same in the two cases, calculate the ratio of the electromagnetic damping of the coil of the first instrument to that of the needle of the second. Ans. 6330.

97. Induction Coil. — This instrument consists essentially of a primary and a secondary, the latter wound outside the former on a longitudinally laminated iron core. A constant electromotive force $\mathcal{E}_1$ is suddenly inserted in the primary circuit and then suddenly removed. By making the number of turns in the secondary much greater than in the primary a momentary e.m.f. much greater than $\mathcal{E}_1$ may be induced in the former. We have here a special case of coupled circuits, which are discussed in more detail in Chapter XV.

First we shall develop the theory of an ideal induction coil in which the e.m.f. is introduced into and removed from the primary without changing its resistance, that is, without breaking the circuit. Designating the primary by the subscript (1) and the secondary by (2) the equations of the two circuits after the insertion of $\mathcal{E}_1$ are

$$L_1\frac{di_1}{dt} + M\frac{di_2}{dt} + R_1i_1 = \mathcal{E}_1, \tag{97–1}$$

$$L_2\frac{di_2}{dt} + M\frac{di_1}{dt} + R_2i_2 = 0, \tag{97–2}$$

where M is the mutual inductance.

First we can simplify (97–1) by the substitution

$$i_1 = i_1' + \frac{\mathcal{E}_1}{R_1},$$

getting

$$L_1\frac{di_1'}{dt} + M\frac{di_2}{dt} + R_1i_1' = 0. \tag{97–3}$$

If we solve this for the derivative of i_2 and differentiate to get

the second derivative as well, we have

$$M\frac{di_2}{dt} = -\left(L_1\frac{di_1'}{dt} + R_1 i_1'\right),$$

$$M\frac{d^2i_2}{dt^2} = -\left(L_1\frac{d^2i_1'}{dt^2} + R_1\frac{di_1'}{dt}\right).$$

If, now, we differentiate (97–2) the resulting equation involves the first and second derivatives of i_2 only. Substituting the expressions above we eliminate i_2, getting

$$(L_1L_2 - M^2)\frac{d^2i_1'}{dt^2} + (R_1L_2 + R_2L_1)\frac{di_1'}{dt} + R_1R_2 i_1' = 0. \qquad (97\text{–}5)$$

This equation is of the same form as that (89–5) of an isolated circuit containing self-inductance, capacity and resistance. Referring to article 89 we see that

$$i_1' = e^{-\alpha t}(A_1e^{\beta t} + A_2e^{-\beta t}), \qquad (97\text{–}6)$$

where in this case

$$\alpha \equiv \frac{R_1L_2 + R_2L_1}{2(L_1L_2 - M^2)}, \qquad \beta \equiv \frac{\sqrt{(R_1L_2 - R_2L_1)^2 + 4M^2R_1R_2}}{2(L_1L_2 - M^2)}.$$

As the expression under the radical is always positive β is always real and the circuit is aperiodic. As equations (97–2) and (97–3) are of exactly the same form it is clear that

$$i_2 = e^{-\alpha t}(B_1e^{\beta t} + B_2e^{-\beta t}), \qquad (97\text{–}7)$$

where, on account of the symmetry of the expressions for α and β above, the exponents in this equation are identical with those in the equation for i_1'. As $i_2 = 0$ when $t = 0$, $B_1 = -B_2 \equiv -B$ and we may write (97–7)

$$i_2 = -Be^{-\alpha t}(e^{\beta t} - e^{-\beta t}). \qquad (97\text{–}8)$$

Now L_1L_2 is always greater than M^2. Consequently α and β are positive, and as $\alpha > \beta$ both terms in (97–8) approach zero for very large t, the second always being smaller in absolute magnitude than the first. The current in the secondary, then,

starts with the value zero, increases to a maximum, and finally falls off asymptotically with the time.

As $i_1 = 0$ when $t = 0$, $i_1' = -\mathcal{E}_1/R_1$. Therefore

$$A_1 + A_2 = -\frac{\mathcal{E}_1}{R_1}$$

from (97–6). Also we get two relations between A_1, A_2 and B if we substitute (97–6) and (97–8) in (97–2) and equate to zero the coefficients of each of the exponentials. Solving we find

$$B = \frac{M\mathcal{E}_1}{R_1R_2}\left(\frac{\alpha^2 - \beta^2}{2\beta}\right),$$

$$A_1 = -\frac{\mathcal{E}_1}{2\beta R_1}\left\{\alpha + \beta - \frac{L_2}{R_2}(\alpha^2 - \beta^2)\right\},$$

$$A_2 = \frac{\mathcal{E}_1}{2\beta R_1}\left\{\alpha - \beta - \frac{L_2}{R_2}(\alpha^2 - \beta^2)\right\},$$

so we have finally for the two currents

$$i_1 = \frac{\mathcal{E}_1}{R_1}\left[1 - \frac{e^{-\alpha t}}{2\beta}\left\{(\alpha + \beta)e^{\beta t} - (\alpha - \beta)e^{-\beta t} - \frac{L_2}{R_2}(\alpha^2 - \beta^2)(e^{\beta t} - e^{-\beta t})\right\}\right], \quad (97\text{–}9)$$

$$i_2 = -\frac{M\mathcal{E}_1}{R_1R_2}\frac{\alpha^2 - \beta^2}{2\beta}e^{-\alpha t}(e^{\beta t} - e^{-\beta t}). \quad (97\text{–}10)$$

The current in the primary approaches asymptotically the steady value $\mathcal{E}_1/R_1$. As is indicated by the negative sign in (97–10) the current in the secondary has the opposite sense to that in the primary. By setting the derivative of i_2 with respect to t equal to zero we find that it is greatest at the time

$$t_m = \frac{1}{2\beta}\log\frac{\alpha + \beta}{\alpha - \beta}. \quad (97\text{–}11)$$

The equations for the decay of the current are (97–2) and (97–3), i_1' now representing the current in the primary. When $t = 0$, $i_1' = \mathcal{E}_1/R_1$ and $i_2 = 0$, so our initial conditions have been altered only in that the sign of $\mathcal{E}_1$ has been changed. Hence we

can write down the expressions for the currents at once, using (97–6) to get i_1'. Dropping the prime we have

$$i_1 = \frac{\mathcal{E}_1}{R_1}\frac{e^{-\alpha t}}{2\beta}\left\{(\alpha+\beta)e^{\beta t} - (\alpha-\beta)e^{-\beta t} - \frac{L_2}{R_2}(\alpha^2-\beta^2)(e^{\beta t} - e^{-\beta t})\right\}, \quad (97\text{–}12)$$

$$i_2 = \frac{M\mathcal{E}_1}{R_1R_2}\frac{\alpha^2-\beta^2}{2\beta}e^{-\alpha t}(e^{\beta t} - e^{-\beta t}). \quad (97\text{–}13)$$

We notice that as the primary decays from $\mathcal{E}_1/R_1$ to zero the

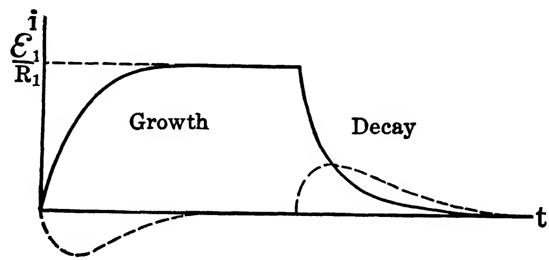


FIG. 211

current in the secondary has the same sense as that in the primary. The secondary current assumes its greatest value at the same time (97–11) as in the case of the growth of current in the primary. The curves in Fig. 211 illustrate the phenomenon under consideration, the solid curve representing the current in the primary and the broken curve that in the secondary.

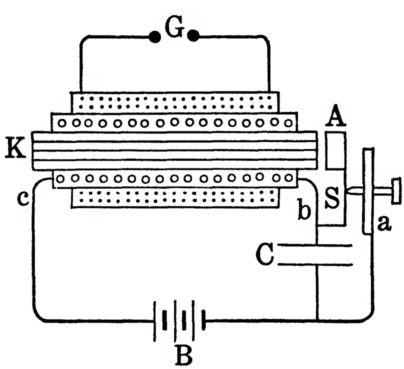


FIG. 212

The actual induction coil is illustrated schematically in Fig. 212. The primary circuit contains the battery B and the condenser C which is placed in 'parallel with the mechanical make and break. The latter consists of a soft iron armature A on the end of the spring S. In growing, the

primary current follows the path $BabcB$, but after it has reached a certain magnitude the soft iron core K of the coil becomes sufficiently magnetized to draw A toward it and break the circuit at S. The primary current now is constrained to follow the path $BCbcB$, charging the condenser. In this way oscillations are set up which last until the current has decayed to the point where K loses enough of its magnetization to release A and allow contact at S to be renewed.

Generally the growth of the primary current is not sufficiently rapid to produce an e.m.f. in the secondary great enough to cause a spark to jump the gap G. The decay, on the other hand, is much more rapid, and the rate of change of current in the primary and therefore the induced e.m.f. in the secondary is increased by the oscillations due to the presence of the condenser.

In order to investigate the essential characteristics of the circuits after the break of the primary current without becoming involved in too lengthy an analysis we shall neglect the resistance of the primary. The resistance of the secondary, on the other hand, is considerable and must be taken into account. We have then for the equation of the primary after the break

$$L_1 \frac{di_1}{dt} + M \frac{di_2}{dt} + \frac{q_1}{C_1} = \mathcal{E}_1, \qquad (97\text{–}14)$$

which can be expressed in terms of the currents alone if we differentiate with respect to the time and put i_1 for $\frac{dq_1}{dt}$. Then we have for the two circuits

$$L_1 \frac{d^2 i_1}{dt^2} + M \frac{d^2 i_2}{dt^2} + \frac{1}{C_1} i_1 = 0, \qquad (97\text{–}15)$$

$$L_2 \frac{di_2}{dt} + M \frac{di_1}{dt} + R_2 i_2 = 0, \qquad (97\text{–}16)$$

since $\mathcal{E}_1$ in (97–14) is the constant e.m.f. of the battery B.

Differentiating (97–15) again with respect to the time,

$$L_1 \frac{d^3 i_1}{dt^3} + M \frac{d^3 i_2}{dt^3} + \frac{1}{C_1} \frac{di_1}{dt} = 0,$$

and from (97–16),

$$M\frac{di_1}{dt} = -L_2\frac{di_2}{dt} - R_2 i_2,$$

$$M\frac{d^3 i_1}{dt^3} = -L_2\frac{d^3 i_2}{dt^3} - R_2\frac{d^2 i_2}{dt^2}.$$

Using the last pair to eliminate i_1 from the previous equation,

$$(L_1L_2 - M^2)\frac{d^3 i_2}{dt^3} + R_2L_1\frac{d^2 i_2}{dt^2} + \frac{L_2}{C_1}\frac{di_2}{dt} + \frac{R_2}{C_1}i_2 = 0.$$

The solution is of the form

$$i_2 = Ae^{mt},$$

where m is a root of the equation

$$am^3 + bm^2 + cm + d = 0, \qquad (97\text{–}17)$$

the positive constants a, b, c, d being

$$a \equiv L_1L_2 - M^2, \qquad b \equiv R_2L_1, \qquad c \equiv \frac{L_2}{C_1}, \qquad d \equiv \frac{R_2}{C_1}.$$

Now the cubic (97–17) has one real negative root as the function of m on the left changes sign between $m = -\infty$ and $m = 0$. The magnitudes of the constants involved in the problem under consideration are such that the two remaining roots are conjugate complex quantities, the real part of each being negative. So if we denote the three roots by $-\gamma$, $-\alpha + i\omega_0$, $-\alpha - i\omega_0$,

$$i_2 = A_0e^{-\gamma t} + e^{-\alpha t}(A_1e^{i\omega_0 t} + A_2e^{-i\omega_0 t}).$$

As was shown in article 89 this can be written

$$i_2 = A_0e^{-\gamma t} - Ae^{-\alpha t}\sin(\omega_0 t + \epsilon). \qquad (97\text{–}18)$$

If we count time from the instant of breaking the primary circuit, $i_2 = 0$ when $t = 0$. Therefore $A_0 = A\sin\epsilon$. The current is given by (97–18) as the sum of an exponentially decreasing term and a damped harmonic oscillation. These two components

are represented by broken lines in Fig. 213, their sum being given by the solid line. Actual oscillograms of the current in the secondary of an induction coil conform quite closely to the

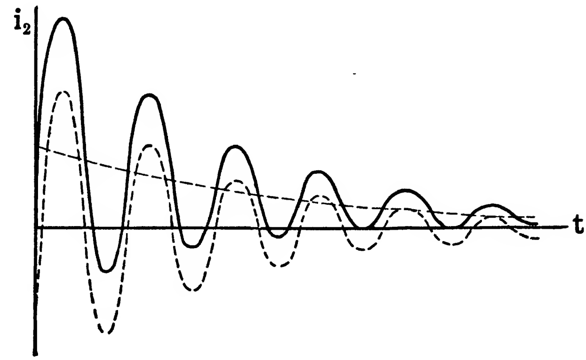


FIG. 213

theoretical curve which we have deduced, when the secondary terminals are connected by a constant high resistance. When they are not so connected and a spark is produced, the resistance of the gap decreases materially after the initial discharge and the current curve is somewhat distorted although it retains the essential features of the graph depicted in the figure.

Problem 97a. Show that the maximum current in the secondary of the ideal induction coil discussed at the beginning of this article is

$$\frac{M\mathcal{E}_1}{R_1R_2}\frac{(\alpha-\beta)^{(\alpha/2\beta)+1}}{(\alpha+\beta)^{(\alpha/2\beta)-1}}.$$

Problem 97b. Show that (97–9) reduces to (87–8) when $M = 0$.

Problem 97c. The gap G (Fig. 212) in the secondary of an induction coil is so wide that no current passes. Find the potential drop across G when the primary current is interrupted. Neglect the resistance in the primary but take account of the capacity of the condenser C.

$$\text{Ans.}\ -\frac{Mi_0}{\sqrt{L_1C_1}}\sin\cdot\ t$$

CHAPTER XI

FLUX MEASUREMENTS

98. Ballistic Galvanometers. — It is often desirable to measure the total quantity of electricity passing through a circuit due to a momentary current. For instance, if a condenser of capacity C charged to a potential difference V is allowed to discharge, a quantity of electricity $Q = CV$ passes through the circuit connecting its plates. By measuring Q we may determine C, if V is known.

A more important case occurs in measurements involving magnetic induction. Consider a coil through which the flux of induction due to external sources changes suddenly from N_1 to N_2. During the short interval in which the flux is changing there is an induced e.m.f. $-\frac{dN}{dt}$ and if the coil forms part of a circuit whose total self-inductance is L and total resistance is R there is a momentary current i given by

$$L\frac{di}{dt} + Ri = -\frac{dN}{dt}.$$

Therefore, integrating over the time during which the current flows, the charge passing through the circuit is

$$Q = \int idt = -\frac{1}{R}\left\{L\int_0^0 di + \int_{N_1}^{N_2} dN\right\}.$$

This reduces to

$$Q = \qquad \frac{\qquad}{R}, \qquad (98\text{-}1)$$

so that determination of flux may be made to depend on measurement of charge.

A *ballistic galvanometer* is one designed to measure the quantity of electricity passing through it in terms of the amplitude of

the mechanical oscillation produced by the momentary current. An ordinary instrument of either the moving needle or the moving coil type, such as were described in article 74, may be used ballistically provided (*a*) it has a period of oscillation long compared to the duration of the momentary current, and (*b*) the damping of the moving element is not too great. On account of the second requirement the damping loop (art. 96) of a D'Arsonval galvanometer must be removed before it is used to measure charge. Also, if the galvanometer is used in a low resistance circuit, it may be necessary to break the circuit immediately after the passage of the charge and before the moving element has suffered an appreciable deflection. Although ordinary galvanometers are usually satisfactory for the measurement of charge as well as of current, special ballistic galvanometers with longer period and small damping are available for precise work.

To investigate the ballistic behavior of a galvanometer let I be the moment of inertia of the moving part and α its angular deflection. Then its equation of motion is

$$I\frac{d^2\alpha}{dt^2} = -f(\alpha) - b'\frac{d\alpha}{dt} + g(\alpha)i, \tag{98–2}$$

where $f(\alpha)$ is the restoring torque, $b'\dfrac{d\alpha}{dt}$ is the total damping torque, and $g(\alpha)i$ is the deflecting torque due to the current i. In the case of the tangent galvanometer $f(\alpha)$ is the torque $MH_e \sin\alpha$ (art. 74) exerted by the control field on the needle, whereas with the astatic or the moving coil galvanometer $f(\alpha)$ is the torque $k\alpha$ due to the suspension. The quantity b', which includes both mechanical and electromagnetic damping, reduces to b (art. 96) when the latter effect is absent.

We consider first the short interval τ during which the momentary current i passes through the instrument. During this time the moving part acquires momentum from the impulse given to it by the current, but, on account of its long period, its deflection $\Delta\alpha$ is only a negligible fraction of that which it acquires

after the current ceases to flow. So, integrating (98–2),

$$I\frac{d\alpha}{dt} = -\int_T f(\alpha)dt - b'\Delta\alpha + \int_T g(\alpha)idt$$
$$= -f(0)\int_T dt - b'\Delta\alpha + g(0)\int_T idt,$$

where we have replaced α by 0 in the functions f and g since the deflection during the short interval of time under consideration is quite negligible. Now $f(0)$ is zero for all types of instrument. Therefore the first term on the right vanishes. Next, the second term is inappreciable on account of the smallness of $\Delta\alpha$. So, as the charge passing through the instrument is

$$Q = \int_T idt,$$

we have for the angular velocity acquired

$$\frac{d\alpha}{dt} = CQ. \tag{98–3}$$

For the tangent galvanometer $g(0) = 2\pi nM/a \equiv MH_e/K_i$ by (74–1), K_i being the galvanometer constant, and

$$C \equiv \frac{g(0)}{I} = \frac{MH_e}{K_iI}. \tag{98–4}$$

For the astatic moving needle galvanometer, on the other hand, $g(0) = 2\pi nM/a \equiv k/K_i$ in accord with (74–7). Hence

$$C \equiv \frac{g(0)}{I} = \frac{k}{K_iI}. \tag{98–5}$$

Finally, for the moving coil type, $g(0) = nAH \equiv k/K_i$ by (74–8), where n is the number of turns in the moving coil, A the area of each and H the field of the fixed magnet. So the constant C is given by (98–5) for this instrument as well.

Next we must consider the motion after the current has ceased to flow. As $i = 0$, (98–2) becomes

$$I\frac{d^2\alpha}{dt^2} + b'\frac{d\alpha}{dt} + f(\alpha) = 0, \tag{98–6}$$

subject to the initial conditions $\alpha = 0$ and $\dfrac{d\alpha}{dt} = CQ$ when $t = 0$. In discussing this part of the motion we must treat the different types of galvanometer separately, since $f(\alpha)$ has different functional forms in the different cases.

Tangent Galvanometer. — In order to arrive at a simple solution we shall suppose that damping is negligible. Then (98–6) becomes

$$I\frac{d^2\alpha}{dt^2} + MH_e \sin\alpha = 0. \qquad (98\text{–}7)$$

Multiplying by $\dfrac{d\alpha}{dt}dt$ and integrating we get

$$\frac{1}{2}I\left(\frac{d\alpha}{dt}\right)^2 = MH_e \cos\alpha + A.$$

Determining the constant of integration A by the initial conditions given in connection with (98–6),

$$C^2Q^2 - \left(\frac{d\alpha}{dt}\right)^2 = 2\frac{MH_e}{I}(1 - \cos\alpha)$$
$$= 4\frac{MH_e}{I}\sin^2\frac{\alpha}{2}.$$

We observe the amplitude α_0 of the oscillation, which is known technically as the *throw* of the galvanometer. Since $\dfrac{d\alpha}{dt} = 0$ when $\alpha = \alpha_0$, the charge Q which has passed through the instrument is given in terms of the throw by the equation

$$Q = \frac{2}{C}\sqrt{\frac{MH_e}{I}}\sin\frac{\alpha_0}{2}$$
$$= 2K_i\sqrt{\frac{I}{MH_e}}\sin\frac{\alpha_0}{2}. \qquad (98\text{–}8)$$

Now for small oscillations of the needle (98–7) becomes

$$I\frac{d^2\alpha}{dt^2} + MH_e\alpha = 0,$$

since we may replace $\sin \alpha$ by α itself. This is simple harmonic motion with the period

$$P_0 = 2\pi \sqrt{\frac{I}{MH_e}}.$$

Therefore we may write (98–8) in the form

$$Q = \frac{P_0}{\pi} K_i \sin \frac{\alpha_0}{2}, \tag{98–9}$$

which becomes

$$Q = K_q \alpha_0, \qquad K_q \equiv \frac{P_0 K_i}{2\pi}, \tag{98–10}$$

if the throw is small. This type of ballistic galvanometer is an absolute instrument since K_i and the period P_0 may be determined by direct measurement. It must be remembered that P_0 is the period for small oscillations; if the amplitude is considerable, the period observed is greater than P_0.

Astatic and Moving Coil Instruments. — In either of these cases (98–6) becomes

$$I \frac{d^2\alpha}{dt^2} + b' \frac{d\alpha}{dt} + k\alpha = 0. \tag{98–11}$$

Dividing by I and putting

$$2l \equiv \frac{b'}{I}, \qquad \kappa^2 \equiv \frac{k}{I},$$

it takes the simpler form

$$\frac{d^2\alpha}{dt^2} + 2l \frac{d\alpha}{dt} + \kappa^2 \alpha = 0. \tag{98–12}$$

This second order differential equation is of precisely the same form as (89–5). Since the damping is small in the present instance we want the periodic solution (89–10),

$$\alpha = \alpha_0 e^{-lt} \sin (\omega_0 t + \epsilon), \qquad \omega_0 \equiv \sqrt{\kappa^2 - l^2};$$

the constants being determined by replacing L, R and C in (89–11) by 1, $2l$ and $1/\kappa^2$ respectively.

To determine α_0 and ϵ we have $\alpha = 0$ when $t = 0$ showing that $\epsilon = 0$, and

$$CQ = \left(\frac{d\alpha}{dt}\right)_{t=0} = \frac{2\pi}{P_0}\alpha_0,$$

since $2\pi/\omega_0$ is the period P_0. Hence

$$\alpha = \alpha_0 e^{-lt} \sin \omega_0 t, \qquad \alpha_0 \equiv \frac{CP_0}{2\pi} Q; \tag{98-13}$$

which is the equation of *damped harmonic motion*, shown in Fig. 214. The curve representing the motion lies between the

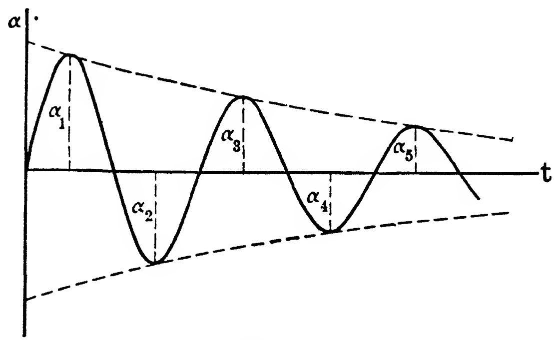


FIG. 214

broken line curves corresponding to $\alpha = \pm\, \alpha_0 e^{-lt}$, being tangent to them when $\sin \omega_0 t = \pm\, 1$.

When the damping is so small that l^2 may be neglected in comparison with κ^2, as is often the case, Q is easily expressed in terms of measurable quantities. For then the period P_0 reduces to $2\pi/\kappa = 2\pi\sqrt{I/k}$ and the time of the first throw α_1 may be taken as that of the first tangency in the figure, that is, $P_0/4$. Replacing C by its value k/K_iI, (98–13) gives

$$Q = \frac{P_0 K_i}{2\pi} e^{(lP_0)/4}\, \alpha_1 \tag{98-14}$$

and it remains only to evaluate the exponential factor.

Since the amplitudes of successive throws are

$$\alpha_1 = \alpha_0 e^{-(lP_0)/4}, \qquad \alpha_2 = \alpha_0 e^{-(3lP_0)/4}, \qquad \alpha_3 = \alpha_0 e^{-(5lP_0)/4}, \cdots$$

each is in the constant ratio $e^{(lP_0)/2}$ to the next. Denoting by λ the logarithm of this quantity,

$$\lambda = \frac{lP_0}{2} = \log\frac{\alpha_1}{\alpha_2} = \log\frac{\alpha_2}{\alpha_3} = \cdots. \tag{98-15}$$

This is one-half the logarithmic decrement as defined in article 89, the half logarithmic decrement being more convenient here.

To obtain λ experimentally it is only necessary to observe any two successive throws. However, when the damping is small, the ratio of successive throws is nearly unity, and a more accurate value is obtained by observing throws separated by several half periods; for evidently we may express λ in the form

$$\lambda = \frac{1}{r}\log\frac{\alpha_n}{\alpha_{n+r}}, \tag{98-16}$$

and the effect of experimental error is reduced by a factor $1/r$.

Returning to (98–14),

$$e^{(lP_0)/4} = e^{\lambda/2} = 1 + \frac{\lambda}{2},$$

as we are neglecting terms in l^2, and hence

$$Q = K_q\left(1 + \frac{\lambda}{2}\right)\alpha_1, \qquad K_q \equiv \frac{P_0 K_i}{2\pi}. \tag{98-17}$$

Since K_i is obtained by passing a steady current through the instrument as explained in article 74, and P_0 and λ may be observed, the ballistic galvanometer is an absolute device, that is, it does not require calibration. However, it is often simpler to pass a known charge through it, from a condenser, for example, and to observe the first throw, thereby obtaining the constant $K_q' \equiv K_q(1 + \lambda/2)$ at once.

When the damping is not small enough to permit the neglect of terms in l^2, the analysis is slightly more involved. This case occurs quite frequently, for instance in the measurement of the magnetic induction B. Here a small closely wound *flux coil* of area A and n turns is connected to the galvanometer and

placed in the field at the point at which B is to be measured, with its plane perpendicular to the lines of induction. The coil is now suddenly moved to a place where the field is zero, or the field at the coil is reduced to zero by removing its source. The charge passing through the galvanometer due to change of flux through the coil is given by (98–1), where $N_1 = BAn$ and $N_2 = 0$, so that

$$B = \frac{RQ}{An}. \tag{98–18}$$

Note that R and Q must be expressed in e.m.u. to give B in gauss.

Now the electromagnetic damping (art. 96) varies inversely with R, in this case the combined resistance of galvanometer and external coil. As R is relatively not very great, the damping, at least for a moving coil instrument, is quite large. It is, of course, possible to open the circuit just as the galvanometer swing begins, but this is often troublesome in practice. A moving needle instrument does not suffer from excessive damping but is much less convenient to use.

To obtain the exact relation between Q and α_1 we return to (98–13). As before $C = k/K_iI$, but now

$$P_0 = \frac{2\pi}{\sqrt{\kappa^2 - l^2}} = \frac{2\pi}{\sqrt{\frac{k}{I} - l^2}},$$

so that

$$\frac{k}{I} = \frac{4\pi^2}{P_0^2} + l^2,$$

and

$$\alpha_0 = \frac{P_0}{2\pi K_i}\left(\frac{4\pi^2}{P_0^2} + l^2\right)Q = \frac{2\pi}{P_0 K_i}\left(1 + \frac{\lambda^2}{\pi^2}\right)Q. \tag{98–19}$$

The times t_m at which the throws $\alpha_1, \alpha_2, \alpha_3, \cdots$ occur are obtained by setting $\frac{d\alpha}{dt}$ equal to zero, which gives

$$\tan \omega_0 t_m = \frac{\omega_0}{l}. \tag{98–20}$$

As the period of this function is π/ω_0, it is evident that successive throws are separated by a half period as before, and that the ratio of any two successive throws is $e^{(lP_0)/2}$. Thus λ is determined exactly as in the previous case. However, the time T of the first throw α_1 is not $P_0/4$ but somewhat earlier, as indicated by (98–20).

Since

$$\sin \omega_0 t_m = \frac{1}{\sqrt{1 + \frac{l^2}{\omega_0^2}}} = \frac{1}{\sqrt{1 + \frac{\lambda^2}{\pi^2}}},$$

we have for the initial throw

$$\alpha_1 = \frac{2\pi}{P_0 K_i}\left(1 + \frac{\lambda^2}{\pi^2}\right)\frac{e^{-lT}}{\sqrt{1 + \frac{\lambda_2}{\pi^2}}} Q,$$

and hence

$$Q = \frac{P_0 K_i}{2\pi}\frac{e^{lT}}{\sqrt{1 + \frac{\lambda^2}{\pi^2}}}\alpha_1. \tag{98–21}$$

All quantities in the coefficient of α_1 are known, for (98–20) may be put in the form

$$lT = \frac{\lambda}{\pi}\arctan\frac{\pi}{\lambda}$$

to determine lT, but usually we write $Q = K_q'\alpha_1$ and determine the constant by calibration.

The expression (98–21) is valid for any magnitude of damping, provided the motion of the galvanometer element is oscillatory. If the damping is made sufficiently great, the motion becomes non-oscillatory, but even in this case it may be shown that the single throw is proportional to the charge. Thus an equation of the form $Q = K_q'\alpha_1$ holds for a ballistic galvanometer under all circumstances.

Problem 98a. A tangent galvanometer having a period of 8 sec deflects 15° when a steady current of 0.1 amp passes through it. What charge will produce a throw of 30° when the instrument is used ballistically, assuming negligible damping? Ans. 0.246 coulomb.

Problem 98b. A secondary of m turns is wound around the middle of a long air core solenoid of n_1 turns per unit length and cross-section A. The secondary is connected to an astatic ballistic galvanometer, the resistance of the circuit being R. Find the quantity of electricity passing through the galvanometer when the current i in the primary is reversed, expressing all electrical quantities in practical units. If the throws are $\alpha_1 = 10$ scale divisions, $\alpha_2 = 9.5$ scale divisions, etc., and $i = 10$ amp, $R = 500$ ohm, $m = 100$ turns, $n_1 = 10$ turn/cm, $A = 20$ cm^2, find the ballistic constant K_q of the galvanometer in coulombs per scale division.

Ans. $Q = \dfrac{8\pi m n_1 A i}{(10)^9 R}$, $\quad K_q = 0.980(10)^{-6}$ coulomb/scale div.

Problem 98c. The coil of a moving coil ballistic galvanometer has a period of 10 sec for free oscillation, and throws $\alpha_1 = 12$ scale divisions, $\alpha_2 = 11$ scale divisions, etc. A steady current of $(10)^{-4}$ amp produces a deflection of 10 scale divisions. Find the capacity of a condenser which, when charged to 100 volts and then discharged through the galvanometer, produces an initial throw of 8 scale divisions. Ans. 1.33 microfarad.

99. The Grassot Fluxmeter. — This instrument is essentially a moving coil galvanometer in which both mechanical damping and restoring torque are made negligibly small, connected by flexible leads to an external flux coil, sometimes called a *search coil*, similar to that used with an ordinary ballistic galvanometer (p. 404).

Let us suppose that the coil is moved from a position where the magnetic induction is B and the flux linkage, that is, the total flux passing through the circuit of the coil, is $N_0 = BAn$ to a position where the flux linkage is zero. It is immaterial whether the motion is sudden or not. The induced e.m.f. gives rise to a current and hence to a torque on the moving element of the fluxmeter. This torque is of the form ci (art. 74) where c is the flux cut by the moving coil per unit angular deflection. The equation of motion for the coil is therefore

$$I\frac{d^2\alpha}{dt^2} = ci, \qquad (99\text{–}1)$$

if I is its moment of inertia.

In order to express i in terms of N it is necessary to write the

electrical equation of the circuit composed of the external coil and the moving element. Let L be the total inductance of this circuit and R the total resistance. Then

$$L\frac{di}{dt} + Ri = -\frac{dN}{dt} - c\frac{d\alpha}{dt}, \tag{99-2}$$

the last term representing the e.m.f. induced in the fluxmeter coil as it turns in the field of the fixed magnet. Eliminating the term in i between (99–1) and (99–2) gives

$$I\frac{d^2\alpha}{dt^2} = \frac{c}{R}\left\{-\frac{dN}{dt} - c\frac{d\alpha}{dt} - L\frac{di}{dt}\right\}.$$

Let us integrate this equation over a period of time which includes the entire action, that is, over a period at the beginning and end of which both i and $\frac{d\alpha}{dt}$ are zero. Thus,

$$\int_0^0 Id\left(\frac{d\alpha}{dt}\right) = \frac{c}{R}\left\{-\int_{N_0}^0 dN - \int_{\alpha_1}^{\alpha_2} cd\alpha - \int_0^0 Ldi\right\},$$

or

$$0 = \frac{c}{R}\{N_0 - c(\alpha_2 - \alpha_1) - 0\}. \tag{99-3}$$

As the moving element has no restoring torque and therefore no zero position, the deflection interval $\alpha_2 - \alpha_1$ may fall anywhere on the scale with which the instrument is provided. Putting $\alpha_0 \equiv \alpha_2 - \alpha_1$, (99–3) gives

$$N_0 = c\alpha_0, \tag{99-4}$$

and

$$B = C\alpha_0, \qquad C \equiv \frac{c}{An}. \tag{99-5}$$

The constant c must be obtained by calibration. In practice the scale is usually arranged to give N_0 directly.

100. Magnetic Standards. — Several methods are employed for the calibration of flux measuring apparatus. When calibration is to be made directly in terms of the induction B, a standard field is necessary. This is obtained by means of a long air core

solenoid the field at the center of which is very closely $4\pi n_1 i$ in e.m.u. according to (70–5), where n_1 is the number of turns per unit length and i is the current. By tapering the ends of the solenoid or by using auxiliary windings at the ends the end effect can be much reduced and the field made essentially uniform over a considerable region on either side of the center.

Often it is only necessary to calibrate in terms of flux. In this case a standard mutual inductance may be used. Such an inductance usually consists of a standard solenoid of the sort described above with a small closely wound secondary coil about its central portion. If there are m turns in this coil and the cross-section of the solenoid is A, a current i in the latter evidently produces a total flux linkage equal to $4\pi m n_1 A i$. Therefore a known change in current produces a known change in flux. In calibrating a Grassot fluxmeter the search coil may be replaced by the secondary of the mutual inductance as the deflection does not depend in any way on the resistance in the circuit. In the case of a ballistic galvanometer provided with a flux coil, however, both the damping of the instrument and the charge passing through it corresponding to a given change of flux depend on the resistance. Hence the secondary coil of the mutual inductance is usually permanently included in the circuit of the galvanometer and its external coil, thereby assuring the use of the galvanometer under the exact conditions of calibration. An example of this procedure appears in the next article.

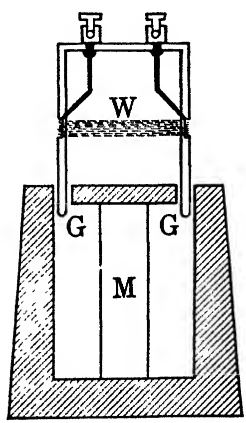


Fig. 215

It is sometimes useful to have a standard source of flux which does not depend on a current. The *Hibbert magnetic standard*, shown in section in Fig. 215, is of this type. A permanent magnet M is set in a cylindrical iron yoke which provides a continuous path for the magnetic flux with the exception of a narrow circular gap G, G. A thin

brass cylinder, carrying a winding W, is arranged to drop through the gap, cutting the flux as it does so. The total flux linkage is found by multiplying the number of turns in the winding by the flux across the gap, which is usually of the order of 20,000 tubes of induction. By using cylinders with different windings different amounts of flux linkage are obtained.

The device just described is, of course, a secondary standard, as the flux across the gap must be determined originally by comparison with an absolute standard such as the mutual inductance described above. It is, however, compact and very convenient, especially in cases where a calibration must be checked frequently.

101. Magnetic Properties of Materials. — The behavior of almost all electrical machinery and apparatus depends on the magnetic properties of some ferromagnetic material. It is therefore very important to be able to determine these properties experimentally. There are several types of measurement which may be made. The most significant of these is the determination of the hysteresis loop, described in article 36. As explained in that article the value of the intensity of magnetization I in a ferromagnetic material depends not only on H but also on the magnetic history of the given specimen. Unless this history has a definite character there is in general no definite relation between I and H. The same is true of B and H since $B = H + 4\pi I$. We find it more convenient to deal with B than with I since B is usually the quantity measured. A repeated cyclic variation of H between limiting values H_m and $-H_m$ provides a definite history and therefore establishes a definite relation—the hysteresis loop—between the variables. A family of loops for sheet iron corresponding to different values of H_m is shown in Fig. 216.

Suppose now we calculate the energy expended in a complete cycle, that is, the work done in carrying a specimen of magnetic material around a given hysteresis loop. As magnetizing fields are usually produced electromagnetically it is interesting to make the calculation from that point of view. For convenience

let the specimen be in the form of a long rod of unit cross-section. If n_1 turns of wire are wound on it per unit length a current i will give rise to a magnetizing field $H = 4\pi n_1 i$. A cyclic variation

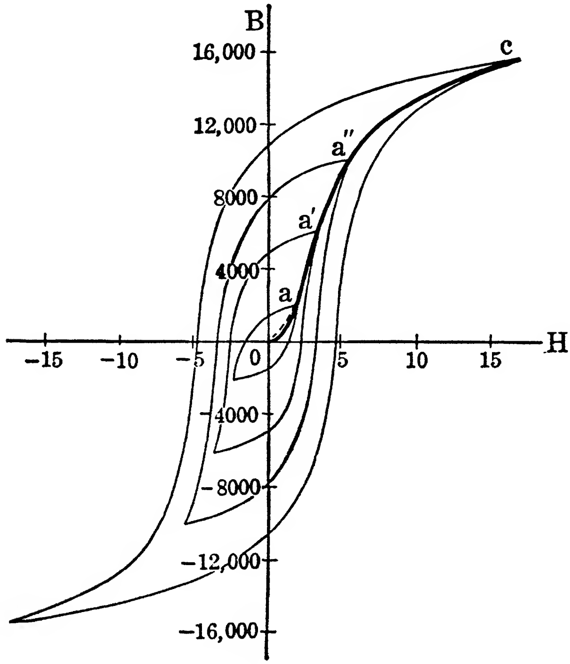


FIG. 216

of i between $i_m = H_m/4\pi n_1$ and $-i_m$ causes the desired variation in H.

Now the variation of current induces a back e.m.f. whose magnitude per unit length of winding is $n_1 \dfrac{dB}{dt}$ by (85–1), since the cross-section of the rod is unity. The work done by the current against this e.m.f. is evidently the magnetic work we are seeking to calculate. Thus for a time dt the work per unit volume of material is

$$dw = n_1 \frac{dB}{dt} idt = n_1 idB = \frac{1}{4\pi} HdB,$$

and for a complete cycle the work per unit volume is

$$w = \frac{1}{4\pi} \oint HdB. \tag{101-1}$$

Since $B = H + 4\pi I$ we may write

$$w = \frac{1}{4\pi} \oint HdH + \oint HdI = \oint HdI, \tag{101-2}$$

for the integral of HdH around a closed path is zero. Equation (101–2) agrees with article 46 where we saw that the work done per unit volume in increasing the magnetization from I to $I + dI$ is HdI. We can, of course, use the results of article 46 to deduce (101–1) instead of the method employed above, if we choose.

Returning to (101–1) we observe that the loop integral represents exactly the area A_{BH} of the $B \sim H$ hysteresis loop, expressed in proper $B \sim H$ units. Thus,

$$w = \frac{1}{4\pi} A_{BH}. \tag{101-3}$$

When a hysteresis loop is plotted from experimental data its area is ordinarily measured by a planimeter or some such device in square centimeters. The value thus obtained must be multiplied by the proper *scale factors*, that is, the number of B units per cm and the number of H units per cm, to obtain A_{BH}. Using e.m.u., w is expressed in erg/cm^3 per cycle.

As the energy represented by (101–3) is dissipated in the form of heat, the hysteresis loop is of basic importance in the choice of materials for electrical machinery of all sorts from an economic as well as from a technical point of view. Care must be taken to distinguish between energy losses due to hysteresis and those due to eddy currents (art. 95), which may be minimized by lamination. A rough idea of the magnitude of hysteresis losses may be obtained from an empirical formula of Steinmetz,

$$w = \eta B_m^{1.6}, \tag{101-4}$$

where B_m is the maximum induction reached in the loop and η,

the *Steinmetz coefficient*, is a constant varying from about 0.001 for soft iron to 0.025 for hard steel.

To illustrate the manner in which the magnetic characteristics of materials are revealed by their hysteresis loops, several loops

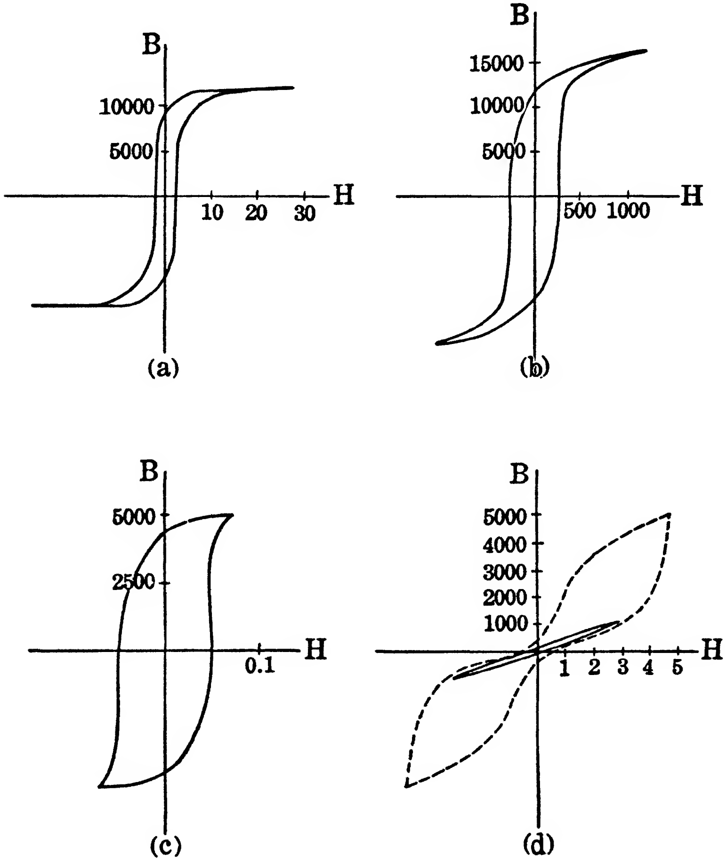


FIG. 217

are shown in Fig. 217. The first loop (*a*) is for soft iron and shows a large saturation value without excessive hysteresis loss. The next loop (*b*) applies to a specimen of cobalt steel. Because of the high retentiveness and large coercive force this material is ideal for permanent magnets but it is not at all suitable for

any cyclic use. Loop (*c*) is for an iron-nickel alloy known as *permalloy*. Its chief characteristic is the high degree of magnetization produced by a very small magnetizing force. Loop (*d*) for *perminvar*, an iron-nickel-cobalt alloy, is of particular interest. For magnetizing fields not greater than about three gauss the loop is very nearly a straight line as indicated by the solid line in the figure, that is, the material behaves like a paramagnetic substance with no hysteresis loss and constant permeability. For greater fields this characteristic disappears and the loop takes the form indicated by the broken line.

Besides the hysteresis loop, another curve of importance is the *normal magnetization curve*. This is the locus of loop tips, shown by $aa'a''c$ in Fig. 216. It is a fictitious curve since the material cannot be carried along it magnetically. However, it is very useful as it shows the maximum induction obtainable with any alternating magnetic force. It should be remarked here that although a general relation for the permeability of the form $\mu = B/H$ can hold accurately only for paramagnetic or diamagnetic substances (art. 43), nevertheless such a relation may be applied to ferromagnetic substances in cases like this, where there is a definite connection between B and H. The *normal permeability* of a ferromagnetic substance is defined as that given by the normal magnetization curve.

The *initial magnetization curve*, which is the actual curve followed by the material when it starts from a state of complete demagnetization, coincides with the normal curve except in the vicinity of the origin, where it is indicated by a broken line on the diagram. Demagnetization is best obtained by placing the specimen in an alternating field of continuously diminishing amplitude. This produces a hysteresis loop that becomes smaller and smaller until it finally vanishes, leaving the material demagnetized. Demagnetization is also obtained by heating, ferromagnetic materials becoming practically non-magnetic at sufficiently high temperatures. However, heating may cause permanent changes in the magnetic characteristics of a substance, especially in the case of an alloy.

The *incremental permeability* μ_i is the effective permeability of a material to a small alternating field superposed on a larger constant one. Thus, if a small change ΔH in the magnetizing field H causes a change ΔB in the induction, $\mu_i \equiv \dfrac{\Delta B}{\Delta H}$. This case occurs frequently in communication apparatus where coils carry d.c. and a.c. currents simultaneously.

Problem 101a. The core of a generator armature is made of iron whose hysteresis loop under operating conditions has an area of $2(10)^5$ $B \sim H$ units. The core is cylindrical, having a length of 40 cm and diameter of 20 cm. If it rotates at 1200 r.p.m. find the rate at which heat is developed in it. Ans. $5.7(10)^3$ cal/min.

Problem 101b. A solenoid 100 cm long and 5 cm in diameter has 500 turns of wire. The solenoid is filled with an iron core whose magnetic characteristics are shown in Fig. 216. Starting with the core demagnetized a current of 0.5 amp is allowed to flow through the winding, a number of reversals being made to establish a cyclic state. Then the current is increased to 1.0 amp and the process repeated. Find the flux of induction in the core in both cases, neglecting end effects. Ans. $0.98(10)^5$ tubes, $2.12(10)^5$ tubes.

102. Measurement of Magnetic Properties. — There are several methods by which the magnetic properties discussed in the previous article may be measured. Description of the more common follows.

Ring Method. — A ring, usually of rectangular cross-section, is cut from the material whose characteristics are to be determined. Two windings are placed on it, a magnetizing coil of n_1 turns per unit length measured at the mean circumference, and a flux coil of m turns total. As the dimensions of the cross-section A are made small in comparison with the mean radius and there are no end effects, a current i in the magnetizing coil produces a very nearly uniform field of magnitude $4\pi n_1 i$ throughout the ring. The corresponding flux of induction is determined with the aid of the flux coil and an associated ballistic galvanometer or fluxmeter. Division by Am gives the induction B itself.

A typical experimental arrangement is shown in Fig. 218.

Current for the magnetizing winding of the ring is supplied by a battery $\mathcal{E}$, the magnitude of the current, which is controlled by an adjustable resistance R_1, being indicated by an ammeter M. The direction of the current is readily reversed by means of a

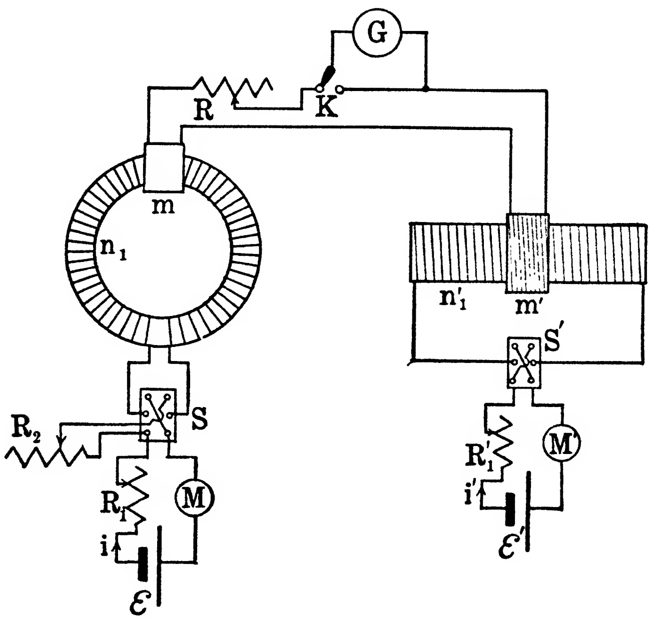


FIG. 218

double-pole double-throw switch S connected as a *reversing switch.* A second resistance R_2 is included with the switch in such a way that it is in series with R_1 when the switch is in one position but not when it is in the other. The flux coil on the ring, with the secondary of a standard mutual inductance and an adjustable resistance R in series, is connected to a ballistic galvanometer G. The mutual inductance is used to calibrate the galvanometer, while R serves to control the sensitivity. Deflection of the galvanometer when not desired is prevented by means of a short-circuiting key K.

The exact experimental procedure depends on the type of measurement. To determine a hysteresis loop R_2 is set at zero

and R_1 is adjusted until the magnetizing field $H = 4\pi n_1 i$ has the desired maximum value H_m for the given loop. Then with the aid of S the field is reversed a number of times, perhaps fifteen or twenty, to establish a cyclic state. This is an important matter and it is well to ascertain the number of reversals necessary by comparing flux measurements made after different numbers. After the cyclic state is established further reversals are without effect. R is set to give approximately a full scale throw of the galvanometer when the maximum field is reversed, and finally G is calibrated. Calibration is effected by observing the throw α_1', corresponding to a reversal of the current i' in the primary of the mutual inductance. Then if a throw α_1 corresponds to a change of induction from B_1 to B_2 in the ring,

$$\frac{(B_1 - B_2)Am}{8\pi m' n_1' A' i'} = \frac{\alpha_1}{\alpha'_1},$$

or

$$B_1 - B_2 = \frac{8\pi m' n_1' A' i'}{m A \alpha_1'}\alpha_1, \qquad (102\text{–}1)$$

the primed quantities referring to the calibrating apparatus. The galvanometer should be calibrated at several points along its scale, in case the latter is not quite linear. Also, it must be recalibrated if the sensitivity is increased as is usually necessary in measuring the smaller values of B.

We are now ready to make the actual measurements. Denoting the " up " position of S as plus and the " down " position as minus, we start at plus and observe the galvanometer throw caused by a reversal with $R_2 = 0$. This corresponds to a change of induction equal to $2B_m$, the value of which is given by (102–1). As $H_m = 4\pi n_1 i_m$ (in e.m.u.) we are able to plot the

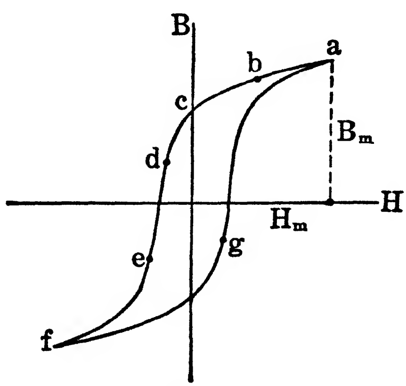


Fig. 219

point *a* (Fig. 219) on the $B \sim H$ diagram. To obtain the next point we return S to plus, bringing the ring back to *a* on the loop, and then increase R_2 until the field is reduced to seven or eight-tenths of its former value, corresponding to point *b*. A reversal of S now gives the change of induction between *b* and *f*, from which the ordinate of *b* is obtained by subtracting B_m. Other points on the portion *ac* of the loop are obtained in the same way. Great care must be taken not to destroy the cyclic state during the measurements. For instance, after the reversal from *b* to *f* just described R_2 must be reduced to zero before S is returned to plus, and then set for the next point; otherwise the ring does not move along the path *fgab* and subsequent observations are incorrect. If the cyclic state is accidentally disturbed it must be reestablished by a number of complete reversals between *a* and *f*.

Points between *c* and *f* are most readily obtained by moving the leads connected to R_1 and M to the other end of S, without changing the polarity. Then R_2 is included in the circuit when S is in the minus position instead of the plus, and, starting at *a* with R_2 correctly set, a reversal carries the ring from *a* to *d*. Subtracting the observed change of induction from B_m, we obtain the ordinate of *d*. For the next point we reduce R_2 to zero, return S to plus and then reset R_2, in order to preserve the cyclic state. As regards the lower branch *fga* of the loop, it is exactly like the upper branch inverted and may be plotted from the data already collected, provided the ring has no permanent magnetic set. As a precaution a few points should be located experimentally for the lower branch even if the ring was originally demagnetized.

Measurements for the normal magnetization curve are relatively simple. R_2 is not used as we require only the location of loop tips. Having determined the coordinates of *a* for a loop in the manner just described, we readjust R_1 to give a different value of H_m and establish a new cyclic state by repeated reversals. The coordinates of the new loop tip are now obtained and so on until enough points are located to delineate a curve from the origin to the region of saturation.

The procedure in measuring incremental permeability is straightforward and will suggest itself to the reader.

Single Bar and Yoke. — In order to avoid the necessity of preparing ring-shaped specimens, Hopkinson devised a heavy yoke of soft iron (Fig. 220) in which a bar of the material to be

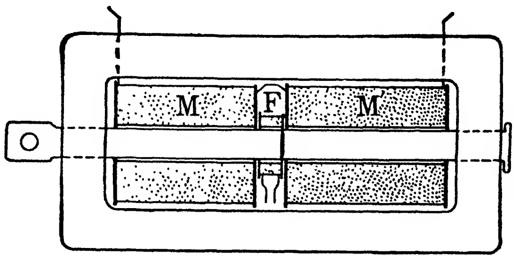


FIG. 220

tested may be placed. Magnetizing coils M, M' and a small flux coil F, all wound on brass bobbins, are slipped over the bar as shown in the figure. The bar is cut at the center and arranged so that one end may be withdrawn a short distance, allowing the flux coil which is attached to a spring to fly out. All the flux in the bar is thereby cut by the coil and B can be measured directly. The corresponding H is calculated from the magnetizing current and the mean turns per unit length of bar. We use the mean turns, that is, the total number of turns divided by the length of bar between the inside surfaces of the yoke, rather than the actual turns per unit length on M and M', because of the small gap in which F lies. The current must be expressed in e.m.u. to give H in gauss.

This method is not highly accurate as there is some leakage of flux where the bar passes through the yoke and at the cut in the center. Moreover, as the yoke does not have infinite permeability, it does not entirely fulfill its purpose of joining the ends of the rod magnetically. Both effects tend to produce demagnetization, that is, to make H less than $4\pi n_1 i$ by a few per cent. We shall see how to calculate a correction for the effect of the yoke in the next article.

The experimental arrangements (Fig. 221) and procedure are very similar to those used with the ring. No auxiliary resistance

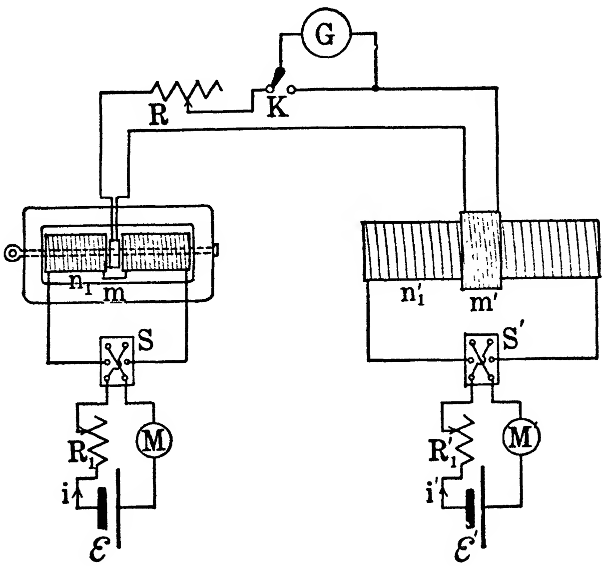


FIG. 221

R_2 is required in this case and the induction is given directly by the formula

$$B = \frac{8\pi m' n_1' A' i'}{m A \alpha_1'} \alpha_1, \qquad (102\text{–}2)$$

the symbols having the same significance as before. One additional precaution must be taken, however. Since each time B is measured the parts of the bar are separated and free poles are exposed, the magnetic condition of the bar is disturbed, and must be restored before another measurement. Thus for a hysteresis loop suppose the preliminary adjustments and galvanometer calibration have been made and the coordinates of the point a (Fig. 219) have been obtained. To locate b the flux coil is returned to the measuring position, the field set at H_m and a number of reversals made to reestablish the cyclic state destroyed by the previous measurement. Then the field is reduced

to the desired value H by adjustment of R_1, and B determined. This process must be repeated for each point, reversals being made at H_m before the field is adjusted to the measuring value. Points between c and f are most easily reached by adjusting the field to a positive value of the same magnitude as the desired negative value, and then throwing S from plus to minus. This fulfills the condition that the bar must always move from a to f on the upper branch of the loop.

The procedure for the normal magnetization curve is exactly the same with the yoke as with the ring, since it is necessary to establish a cyclic state before each measurement in either case.

Double Bar and Yoke. — By modifying Hopkinson's method to use two bars instead of one (Fig. 222), Ewing succeeded in

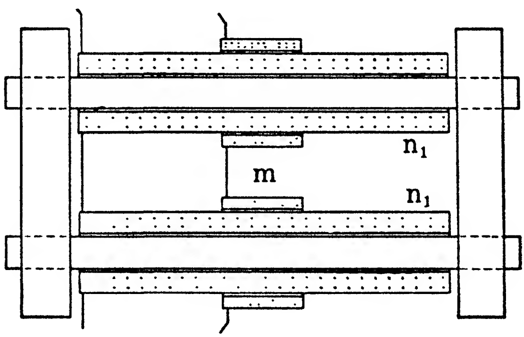


FIG. 222

eliminating the errors due to demagnetizing effects. A bobbin with magnetizing and flux coils is slipped over each bar as shown in the figure. In this case the flux coils are immovable, so all measurements are made in the same manner as when a ring is used.

To understand how the errors are eliminated suppose a set of observations has been made, to determine the normal magnetization curve, for example. The values of B are experimental and the values of H are given by

$$H = H' - h, \qquad (102\text{-}3)$$

where $H' = 4\pi n_1 i'$ and h is the demagnetizing field due to flux leakage and so forth. Let us plot B against H' obtaining curve (1), Fig. 223. The true $B \sim H$ curve (unnumbered) evidently lies to the left of (1) by the amount h, which is a function of B. To find h suppose we replace the two magnetizing coils by another pair each of which has the same number of turns per cm as before, but is one half as long. The yoke blocks are of course moved up to the ends of the coils, so the effective length of the bars is reduced to one-half its former value.

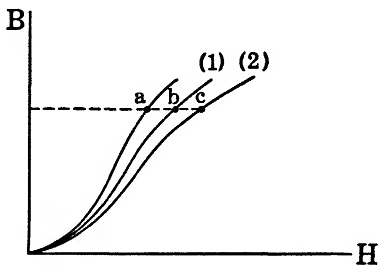


FIG. 223

Repeating the measurements for the $B \sim H$ curve, we have now

$$H = H'' - 2h, \tag{102–4}$$

where $H'' = 4\pi n_1 i''$ and $2h$ is the demagnetizing field. This field is twice as great as before because the same sources of demagnetization become twice as effective with half-length bars, a fact demonstrated analytically in the next article. Plotting B against H'' gives curve (2). Now from (102–3) and (102–4) it appears that curves (1) and (2) are also separated by the amount h. Therefore the true $B \sim H$ curve is simply derived from the other curves by laying off $ab = bc$ for various values of B.

The double bar method is rather laborious and is used in practice only to standardize certain bars for reference purposes.

Magnetometer. — When the quantity of material available for testing is small, the specimen being in the form of a fine wire or thin film, for example, a magnetometer must be used. An ordinary instrument of the sort described in article 47, with the earth's field as a control, will serve, but a double needle or astatic instrument, similar to the astatic galvanometer (art. 74), is preferable. The measurements may be made in a number of ways. An excellent method employing an astatic magnetometer is described by Bozorth (*Review of Scientific Instruments;* May,

1925). Here two identical solenoids (Fig. 224) are symmetrically placed on either side of the magnetometer. The solenoids are connected in series with their fields in the same sense and the whole system carefully balanced until the magnetometer deflection is entirely independent of the solenoids.

The test specimen S is now placed in one solenoid and a very small compensating coil C of the same length as S is placed in the other. Due to the field of the solenoid S becomes a small magnet and produces a deflection of the magnetometer. This deflection is reduced to zero by adjusting the current i' through C until the latter has an equivalent magnetic moment equal to that of S. As this equivalent magnetic moment can be calculated in terms of i' and the dimensions of the coil, the magnetic moment of S becomes known and from this the intensity of magnetization I can be found. H is calculated in terms of the current i in the solenoid by (70–4), taking account of the demagnetizing effect of the ends of the specimen. Finally, using $B = H + 4\pi I$, we can plot B as a function of H.

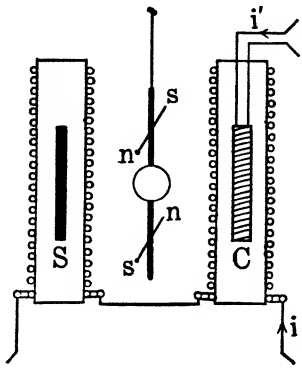


FIG. 224

For details of this and other magnetometer methods, which are rather laborious, the reader should consult some treatise on magnetic measurements.

Permeameters. — For commercial testing a number of complete assemblies of magnetic measuring apparatus have been produced, special care having been taken to insure flexibility, facility of operation, direct indication of B and H, and other convenient characteristics. Such units are called *permeameters.* With one or two exceptions their operation depends on the principle of the yoke.

The best and most widely used permeameter for general testing is the *Fahy simplex permeameter.* It consists of a yoke with several windings (Fig. 225), a control box containing am-

meter, standard inductance, resistances, together with keys and switches to facilitate operation, and a ballistic galvanometer. The construction of the yoke is of especial interest. Its form is shown in the figure. The magnetizing winding M is placed directly on the yoke and does not surround the test specimen S. The latter is clamped in a pair of heavy jaws which are a part of the yoke and passes through a flux coil F, so that B is measured in the usual manner. The determination of H, however, is novel. The sides of the yoke project a short distance beyond F and an air-core flux coil f is placed between them. The field H_f through this coil is not in general equal to the field H in the specimen, which we wish to measure. In fact, H_f is variable both in magnitude and direction, while H is essentially constant. However, the flux linking f is the same as for a constant field equal to H and therefore H can be measured with the aid of f in exactly the same way as B with F.

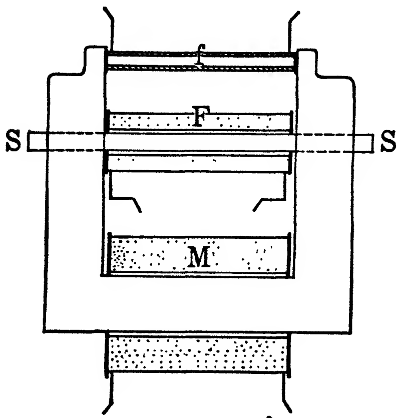


FIG. 225

To demonstrate the equality of flux linkage under the conditions mentioned we observe that the difference of magnetic potential between the ends of f is practically the same as the difference between the ends of the effective portion of S, since the magnetic force and hence the potential drop is very small in the projecting parts of the yoke. Thus if l is the distance between the sides of the yoke, and α is the angle which H_f makes at any point with the axis of f,

$$\int_0^l H_f \cos \alpha dx = Hl, \qquad (102\text{–}5)$$

where dx is an element of length parallel to the axis. Now the

flux linking f is

$$N_f = \int_0^l n_1 dx \int_0^A H_f \cos \alpha ds,$$

n_1 being the turns per cm and A the cross-section, of which ds is an element. Changing the order of integration and using (102–5) gives

$$N_f = HAn_1l, \tag{102–6}$$

which is exactly the flux linkage corresponding to a constant field equal to H, as stated.

The general experimental procedure is the same as in other yoke methods.

103. The Magnetic Circuit. — Having determined the magnetic properties of materials in general, it remains to translate these properties into the production of flux in electromagnetic apparatus and machinery, for all electromagnetic reactions depend in one way or another on flux. We may attack the problem in the same way as in articles 44 and 45, that is, by attempting to find a function specifying the field which satisfies the assigned boundary conditions. This is often difficult, however. Moreover, we are primarily interested in the amount of flux rather than in its exact distribution. The fact that the tubes of induction which comprise the flux are confined to a definite circuital path, composed largely or entirely of ferromagnetic materials, in almost all practical cases, suggests the point of view of a *magnetic circuit.*

To establish the necessary ideas consider the magnetized ring described in the previous article (p. 415). If A is the cross-section of the ring, l the mean distance around it, and n the total number of turns in the magnetizing winding, the flux in the ring, that is, in the magnetic circuit, is

$$N_c = BA = 4\pi\mu \left(\frac{n}{l}\right) iA,$$

or

$$N_c = \frac{4\pi ni}{\dfrac{l}{\mu A}}. \tag{103-1}$$

The flux, being circuital, is the same at any point around the ring. Note that the flux in the magnetic circuit is not the same thing as the flux linking the electric circuit, the latter being n times the former in this case.

Now by Ampère's law $4\pi ni$ is the work done in carrying a unit pole around the magnetic circuit, that is, the magnetomotive force (m.m.f.) defined in article 72. Evidently $l/\mu A$ measures the resistance of the circuit to the production of flux. If we call this magnetic resistance the *reluctance* and write (103–1) in the form

$$\text{flux} = \frac{\text{m.m.f.}}{\text{reluctance}},$$

we have an equation of exactly the same form as Ohm's law, namely,

$$\text{current} = \frac{\text{e.m.f.}}{\text{resistance}}.$$

The magnetic circuit is therefore analogous to an electric circuit, flux corresponding to current, m.m.f. to e.m.f. and reluctance to resistance. The analogy is even more close, for the resistance of a conductor whose length is l, cross-section A and conductivity σ, is $l/\sigma A$, which compares exactly with $l/\mu A$. Thus μ corresponds to σ.

Equation (103–1) deduced from the uniform ring evidently applies with very little error to a ferromagnetic circuit of any shape in which μ and A are constant, even if the magnetizing winding is not uniformly distributed, for the flux is still almost entirely confined to the high permeability material of which the circuit is composed. We may easily generalize (103–1) to apply to a circuit of variable cross-section and permeability, including air gaps. Carrying a unit pole around such a circuit,

$$\oint H dl = 4\pi ni \tag{103-2}$$

by Ampère's law. Now $H = B/\mu = N_c/\mu A$, and, since N_c is constant around the circuit except for slight leakage, (103–2) becomes

$$N_c \oint \frac{dl}{\mu A} = 4\pi n i,$$

so that

$$N_c = \frac{4\pi n i}{\oint \frac{dl}{\mu A}}. \qquad (103\text{–}3)$$

In practice a circuit usually consists of several distinct parts in each of which μ and A are constant. Then

$$\oint \frac{dl}{\mu A} = \frac{l_1}{\mu_1 A_1} + \frac{l_2}{\mu_2 A_2} + \cdots.$$

This indicates that the reluctance of a magnetic circuit is the sum of the reluctances of its parts, just as the resistance of an electric circuit is the sum of its component resistances. In fact, since the basic equations of the two circuits are alike, all the properties which depend on these equations must be the same. Thus parallel reluctances combine in the same way as parallel resistances (art. 52) and flux divides between the former in the same way as current divides between the latter.

Neglecting leakage of flux through the air evidently introduces an error of the same sort as neglecting leakage of current from an electric circuit which is placed in a slightly conducting solution. This error is not very great except perhaps at an air gap, where, unless the gap is very small compared to the linear dimensions of the circuit at that point, there is some spreading of flux. In any case leakage factors are known for the common forms of magnetic circuit, and may be used if necessary.

As an example of the utility of magnetic circuit analysis let us find the effect of cutting a narrow gap of width a in the uniform ring discussed above. Using the same symbols as before the reluctance is now

$$\frac{l - a}{\mu A} + \frac{a}{A} = \frac{l + a(\mu - 1)}{\mu A}.$$

Comparing this with the former value $l/\mu A$ we see that the gap reduces the original flux by a factor $\dfrac{l}{l + a(\mu - 1)}$. For instance, with $\mu - 1 = 500$, a gap as small as $l/500$ reduces the flux by one-half, which shows the great effect of even very small gaps in a magnetic circuit.

Another useful application of (103–3) may be made to yokes used in magnetic measurements (art. 102). Denoting total length of path, cross-section and permeability by l, A and μ respectively for the test material, and by l', A' and μ' for the yoke, we have

$$N_c = \frac{4\pi n i}{\dfrac{l}{\mu A} + \dfrac{l'}{\mu' A'} + \dfrac{a}{A}}, \tag{103–4}$$

where a is the total equivalent air gap in the circuit. Now, as we have seen, B is measured in the determination of magnetic characteristics but H must be calculated. Since $H = N_c/\mu A$ equation (103–4) becomes

$$H = \frac{4\pi n_1 i}{1 + \dfrac{l'\mu A}{l\mu' A'} + \mu\dfrac{a}{l}},$$

n_1 being the mean number of turns in the magnetizing winding per unit length of the test bar. As the last two terms in the denominator are made as small as possible in practice, the equation may be written

$$H = 4\pi n_1 i\left(1 - \frac{l'\mu A}{l\mu' A'} - \mu\frac{a}{l}\right). \tag{103–5}$$

In applying this to the single bar and yoke (p. 419) it is customary to neglect the last term in the parentheses, as it is usually smaller than the second term and somewhat indeterminate. The field correction factor, due to the yoke, is then $(1 - l'\mu A/l\mu' A')$, which may differ from unity by several per cent. In the case of the double bar and yoke (p. 421) $\mu a/l$ is not neglected, as it is a part of the demagnetizing field whose effect is eliminated. This

field is specified by the last two terms in (103–5). It is seen to be inversely proportional to l, as was previously assumed.

Problem 103a. Given a U-shaped electromagnet with n turns of wire in the magnetizing winding. Suppose it has square cross-section of area A, length l and pole separation d. The core has a permeability μ. Using (43–12) show that the force in dynes with which it holds a bar of the same cross-section and material against its poles is approximately

$$\frac{2\pi\mu^2 n^2 i^2}{(l+d)^2},$$

i being the magnetizing current in e.m.u.

Problem 103b. On the basis of the preceding problem show how an electromagnet can be made to serve as a permeameter.

CHAPTER XII

ABSOLUTE STANDARDS AND UNITS

104. Absolute Measurement of Current. — The value of a current in electromagnetic units is determined by the magnetic action between the current and a magnet or a second current. The tangent galvanometer described in article 74 may be used to measure the current passing through it in terms of the dimensions of the instrument and the control field H_e, the latter being determined in absolute units by means of the magnetometer described in article 47. This method, however, is not susceptible of very high precision. Greater accuracy is obtained by measuring a current in terms of the force between two connected coils through which it passes. Instruments designed with this object in view are known as *electrodynamometers*.

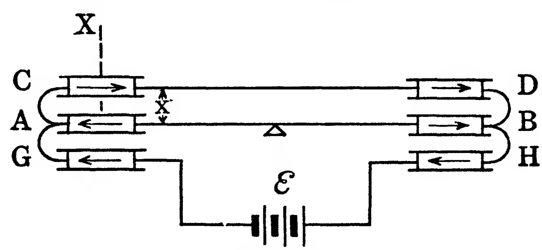


FIG. 226

The *Kelvin current balance* (Fig. 226) is an electrodynamometer consisting of coils A and B attached to the ends of a beam balance, the first lying between the parallel fixed coils C and G and the second between D and H. The coils are connected in series so that the same current i passes through each coil in the sense of the arrow. Consequently A is attracted by G and repelled by C and similarly B is attracted by D and repelled by H. By sliding a rider along the beam of the balance until the

torque due to the forces on A and B is compensated, the latter can be measured in terms of the weight of the rider.

If M is the mutual inductance of A and C, the force between these two coils is

$$i^2 \frac{\partial M}{\partial x}$$

from (93–5), and as the total force F is four times as great,

$$i = \tfrac{1}{2} \sqrt{\frac{F}{\frac{\partial M}{\partial x}}} \tag{104–1}$$

in absolute electromagnetic units. In article 94 we have seen how to calculate the mutual inductance of two parallel coaxial circles in terms of their radii and their separation. Therefore all the quantities on the right are readily determined.

Incidentally it may be noted that, as the force between the coils of an electrodynamometer is proportional to the square of the current, this instrument may be used to measure alternating as well as direct currents. When so used the effective or root-mean-square current is given by formula (104–1). A tangent galvanometer, on the other hand, cannot be employed for this purpose, for the alternations of the torque on the needle occur at such short intervals of time as compared to its period of vibration that no deflection at all is observed. The same statement applies to the D'Arsonval galvanometer.

If, now, a silver voltameter is connected in series with the current balance the mass of silver deposited on the cathode by a current of one ampere flowing for one second can be measured. Early determinations gave 0.0011180 for the electrochemical equivalent of silver, and the International Congress of 1893 adopted as the definition of the *international ampere* that current which will deposit 0.0011180 gm of silver in one second. Later determinations show that the *absolute ampere* (1/10 of the e.m.u. of current) deposits 0.00111805 gm of silver per sec on the cathode of a silver voltameter. Nevertheless the legal standard has not been changed, so the international ampere and coulomb are short

of the corresponding absolute units by a few parts in 100,000. The relation between the international and absolute units of current is

$$1 \text{ int. amp} = (0.99995 \pm 0.00005) \text{ abs. amp.}$$

The discrepancy is quite trivial even for the most precise measurements.

An electrodynamometer designed by Ayrton and Jones for absolute measurements is capable of even greater precision than the Kelvin current balance. This instrument (Fig. 227), considered to be accurate to 1 part in 50,000, consists of two coils *A* and *B* wound on marble cylinders suspended from the ends of a beam balance. The suspended coils move inside the coaxial fixed coils *C* and *D*. All four coils are connected in series in such a way that the force on one of the suspended coils is downward and that on the other upward. The torque due to the current is balanced by a weight in much the same way as with the Kelvin balance.

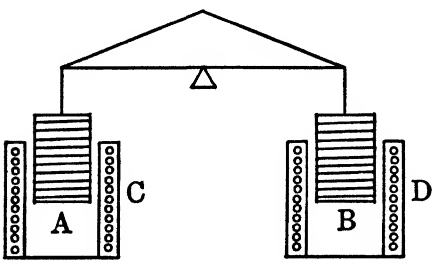


FIG. 227

Once the electrochemical equivalent of silver has been accurately determined, a given current can be measured in absolute units more simply by observing the rate at which it deposits silver in a voltameter than by the use of an electrodynamometer. Furthermore, the calibration curve of a current measuring instrument which is not absolute may be obtained by connecting it in series with a silver voltameter and passing steady currents of various magnitudes through the two, measuring the current in each instance by the rate at which silver is deposited on the cathode of the voltameter.

105. The Absolute Ohm. — The measurement of the resistance of a standard coil of wire in absolute units is of prime im-

portance, since, once a resistance has been so determined, the value in absolute units of an electromotive force may be obtained from Ohm's law in terms of resistance and current. Moreover, any other resistance may be accurately measured in terms of the standard by the use of a Wheatstone bridge.

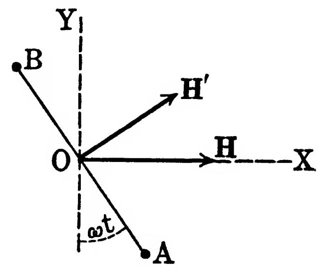


FIG. 228

One method of measuring the resistance of a coil AB of radius a (Fig. 228) in absolute electromagnetic units consists in rotating the coil about a diameter in a uniform field H. If the coil has n closely wound turns, and we take the positive normal to the circuit in the direction of the arrow H', the induced e.m.f. is

$$\mathcal{E} = -\frac{dN}{dt} = \pi n a^2 \omega H \sin \omega t$$

when the coil is rotated about the Z axis with angular velocity ω. This e.m.f. induces a current

$$i = \frac{\pi n a^2 \omega H}{\sqrt{R^2 + L^2\omega^2}} \sin (\omega t - \phi), \qquad \tan \phi \equiv \frac{L\omega}{R},$$

in accord with (87–12). This current in turn gives rise to a magnetic field

$$H' = \frac{2\pi n i}{a} = \frac{2\pi^2 n^2 a \omega H}{\sqrt{R^2 + L^2\omega^2}} \sin (\omega t - \phi)$$

at the center of the coil.

Now if a magnetic needle is suspended at the center O of the rotating coil it will point in the direction of the resultant mean field. The mean values of the components of H' along X and Y are

$$H_x' = \frac{2\pi^2 n^2 a \omega H}{\sqrt{R^2 + L^2\omega^2}} \overline{\sin (\omega t - \phi) \cos \omega t} = -\frac{\pi^2 n^2 a \omega H}{\sqrt{R^2 + L^2\omega^2}} \sin \phi,$$

$$H_y' = \frac{2\pi^2 n^2 a \omega H}{\sqrt{R^2 + L^2\omega^2}} \overline{\sin (\omega t - \phi) \sin \omega t} = \frac{\pi^2 n^2 a \omega H}{\sqrt{R^2 + L^2\omega^2}} \cos \phi,$$

where

$$\sin\phi \equiv \frac{L\omega}{\sqrt{R^2 + L^2\omega^2}}, \qquad \cos\phi \equiv \frac{R}{\sqrt{R^2 + L^2\omega^2}}.$$

Therefore the deflection of the needle from the X axis due to the rotation of the coil is given by

$$\tan\alpha = \frac{H_y'}{H + H_x'} = \frac{\pi^2 n^2 a\omega R}{R^2 + (L^2 - \pi^2 n^2 aL)\omega^2}.$$

For the speeds of rotation ordinarily used the second term in the denominator is negligible compared to the first. Consequently

$$R = \pi^2 n^2 a\omega \cot\alpha, \tag{105-1}$$

which gives the resistance in terms of easily measurable quantities. A small correction is necessary, however, to take account of the current induced in the coil by the magnetic field of the needle.

Other methods of measuring a resistance in absolute units involve comparison of a resistance with a known inductance. A detailed account of such methods may be found in Glazebrook: *Dictionary of Applied Physics.*

The *international ohm* has been established by the London conference of 1908 as the resistance of a column of 14.4521 gm of mercury of uniform cross-section and 106.300 cm length at 0° C. Under these conditions the cross-section is very nearly 1 mm^2. More recent measurements indicate that the international ohm is larger than the absolute ohm by about 5 parts in 10,000, the relation between the two units being

$$1 \text{ int. ohm} = (1.00051 \pm 0.00002) \text{ abs. ohm}.$$

The discrepancy here is considerably greater than that between the international and the absolute ampere and cannot be neglected in precise measurements.

106. Other Absolute Measurements. — The *international volt* is defined as the electromotive force which gives rise to a current of one international ampere in a circuit whose resistance

is one international ohm. On account mainly of the difference between the international and the absolute ohm, the international volt is larger than the absolute volt, the relation between the two units being

$$1 \text{ int. volt} = (1.00046 \pm 0.00005) \text{ abs. volt.}$$

The Weston cell (art. 61), when made according to prescribed specifications, has very closely the e.m.f. 1.0183 int. volt at 20° C when on open circuit. Other electromotive forces can be compared with this by the potentiometer method of article 56.

We have seen that the capacities of condensers of simple form can be calculated in e.s.u. from their dimensions. The capacity of any other condenser can be compared with that of a standard by the bridge method of article 23. In the next article we shall see how the capacity of a condenser may be measured directly in e.m.u., leading to the ratio of the electromagnetic unit of capacity to the electrostatic unit.

Finally both self and mutual inductances may be calculated from their dimensions in many cases by the methods developed in article 94. Once a standard of inductance has been obtained, any other inductance may be measured in terms of it by the bridge described in article 114.

Problem 106a. Find the relation between the international and the absolute joule. Ans. 1 int. joule = (1.00041 ± 0.00010) abs. joule.

107. Relation between Electrostatic and Electromagnetic Units. — The subject of electrostatics is based on Coulomb's law for the force between two point charges, the e.s.u. of charge being defined from the law of force as that charge which repels an equal like charge at a distance of 1 cm *in vacuo* with a force of one dyne. Denoting the charge as measured in e.s.u. by q_s we have

$$F = \frac{q_s q_s'}{r^2}. \qquad (107\text{–}1)$$

In e.m.u. the unit charge is defined as the charge passing through a cross-section of a circuit in one second when one e.m.u.

of current flows in the circuit, the e.m.u. of current being that defined in article 48. If, now, q_m is the measure of a charge in e.m.u. and q_s the measure of the same charge in e.s.u., we must have

$$q_s = cq_m, \tag{107–2}$$

where c is a constant of proportionality which may have physical dimensions. Coulomb's law of force, then, takes the form

$$F = c^2 \frac{q_m q_m'}{r^2} \tag{107–3}$$

when the charges are expressed in e.m.u. Our present aim is to evaluate the dimensional constant c.

We can accomplish this by measuring the capacity of a condenser in both e.s.u. and e.m.u. Distinguishing the two systems of units by the subscripts s and m we have

$$C_s = \frac{q_s}{V_s}, \qquad C_m = \frac{q_m}{V_m}.$$

As both systems of units are c.g.s., the erg is the unit of energy in each. Therefore the energy of the condenser is

$$U = \tfrac{1}{2} q_s V_s = \tfrac{1}{2} q_m V_m.$$

Eliminating the potential,

$$c = \frac{q_s}{q_m} = \sqrt{\frac{C_s}{C_m}}. \tag{107–4}$$

The capacity of a geometrically simple condenser can be calculated in e.s.u. from its dimensions as in article 18, and the capacity of any other condenser can be measured in terms of a standard by the capacity bridge. The capacity of a condenser in e.m.u. may be determined by charging it to a known potential V_m and then discharging it through a calibrated ballistic galvanometer, determining q_m from the throw of the instrument. Then the ratio of q_m to V_m gives the capacity C_m.

Another method of measuring the capacity C_m of a condenser C (Fig. 229) makes use of a vibrating tuning fork. The prong

T of the fork makes connection alternately with A and with B, charging up the condenser from the battery $\mathcal{E}$ through the galvanometer G when in contact with A, and discharging it when in contact with B. If $\mathcal{E}_m$ is the e.m.f. of the battery in e.m.u., n the number of vibrations per second of the tuning fork and i_m the current indicated by galvanometer,

$$i_m = nC_m\mathcal{E}_m,$$

from which C_m can be obtained.

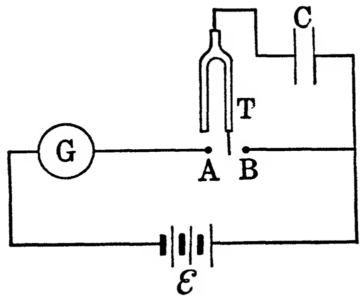


FIG. 229

By taking the square root of the ratio of the numerical value of C_s to that of C_m the numerical part of the ratio of q_s to q_m is found. To find the unit of the latter ratio we have from (107–1)

$$1\,\frac{\text{gm cm}}{\text{sec}^2} = \frac{[q_s]^2}{\text{cm}^2},$$

or

$$[q_s] = \frac{1\ \text{gm}^{1/2}\ \text{cm}^{3/2}}{\text{sec}},$$

if we denote by $[q_s]$ a unit charge in e.s.u.

Similarly from (37–1)

$$[m_m] = \frac{1\ \text{gm}^{1/2}\ \text{cm}^{3/2}}{\text{sec}},$$

and therefore

$$[H_m] = \frac{1\ \text{dyne}}{[m_m]} = \frac{1\ \text{gm}^{1/2}}{\text{cm}^{1/2}\ \text{sec}},$$

$$[i_m] = [H_m]\ \text{cm} = \frac{1\ \text{gm}^{1/2}\ \text{cm}^{1/2}}{\text{sec}},$$

$$[q_m] = [i_m]\ \text{sec} = 1\ \text{gm}^{1/2}\ \text{cm}^{1/2}.$$

Consequently, the ratio of a unit charge in e.s.u. to a unit charge in e.m.u. is

$$\frac{[q_s]}{[q_m]} = 1\,\frac{\text{cm}}{\text{sec}}. \qquad (107\text{–}5)$$

The ratio c of q_s to q_m has been measured by a number of investigators. The latest determination,

$$c = 2.99790(10)^{10} \text{ cm/sec},$$

is that obtained by Rosa and Dorsey of the Bureau of Standards in 1907, as corrected by Curtis in 1929 to accord with the most recent value of the international ohm. On the other hand, Michelson's latest determination of the velocity of light *in vacuo* is $2.99796(10)^{10}$ cm/sec. The close agreement of these two figures leaves no doubt that light itself is an electromagnetic phenomenon and that the two quantities are identical. As the velocity of light can be measured with a somewhat higher degree of precision than the ratio q_s to q_m, the practice at the present time is to take for the latter ratio

$$c = 2.99796(10)^{10} \text{ cm/sec.} \tag{107–6}$$

To a high enough degree of precision for most purposes we may write

$$c = 3(10)^{10} \text{ cm/sec.}$$

This value of the ratio has been used in preparing the table on page xii.

Since $q_s = cq_m$ the electric moment of a dipole is

$$p_s = q_s l = cq_m l = cp_m,$$

and the relation between the polarization in e.s.u. and in e.m.u. is

$$P_s = cP_m. \tag{107–7}$$

The electric intensity is force per unit charge. Therefore

$$q_s E_s = cq_m E_s = q_m E_m,$$

and

$$E_s = \frac{1}{c} E_m. \tag{107–8}$$

In order to preserve the form of Gauss' law (14–4) when expressed in e.m.u. we transform the electric displacement according to the rule

$$D_s = cD_m. \tag{107–9}$$

Then Gauss' law remains

$$\int_s D_m \cos \gamma ds = 4\pi \int_\tau \rho_m d\tau \qquad (107\text{–}10)$$

when D and ρ are specified in e.m.u.

In view of (107–7), (107–8) and (107–9) the relation

$$D_s = E_s + 4\pi P_s = \kappa E_s$$

becomes

$$D_m = \frac{E_m}{c^2} + 4\pi P_m = \frac{\kappa}{c^2} E_m \qquad (107\text{–}11)$$

in e.m.u., the specific inductive capacity κ remaining unchanged since it is a pure ratio.

108. Fundamental Equations. — We are now ready to write down the fundamental equations of electromagnetism in the form they take when we use e.m.u. throughout. We have

$$E_m = c^2 \frac{q_m}{r^2}, \qquad (1\text{m}) \qquad\qquad H_m = \frac{m_m}{r^2}, \qquad (2\text{m})$$

$$\oint \mathbf{H}_m \cdot d\mathbf{l} = 4\pi \int_s \mathbf{j}_m \cdot d\mathbf{s}, \qquad (3\text{m})$$

$$\oint \mathbf{E}_m \cdot d\mathbf{l} = -\int_s \frac{\partial \mathbf{B}_m}{\partial t} \cdot d\mathbf{s}, \qquad (4\text{m})$$

$$D_m = \frac{E_m}{c^2} + 4\pi P_m = \frac{\kappa}{c^2} E_m, \qquad (5\text{m})$$

$$B_m = H_m + 4\pi I_m = \mu H_m. \qquad (6\text{m})$$

Equations (1m) and (2m) are Coulomb's laws for electric charges and magnetic poles respectively, (3m) is Ampère's law, (4m) Faraday's law and (5m) and (6m) are the relations between displacement, electric intensity and polarization on the one hand and induction, magnetic intensity and intensity of magnetization on the other.

To pass to any other system of c.g.s. units we may put $q_m = \alpha q'$, $m_m = \beta m'$, $D_m = \gamma D'$, $B_m = \delta B'$ where we agree to

preserve the equations

$$\left.\begin{aligned} F &= q_m E_m = q' E', \\ F &= m_m H_m = m' H', \\ t &= \frac{q_m}{i_m} = \frac{q'}{i'}. \end{aligned}\right\}(108\text{–}1)$$

Therefore we have $E_m = (1/\alpha)E'$, $H_m = (1/\beta)H'$, $i_m = \alpha i'$, and our equations become

$$E' = \alpha^2 c^2 \frac{q'}{r^2}, \qquad (1) \qquad\qquad H' = \beta^2 \frac{m'}{r^2}, \qquad (2)$$

$$\oint \mathbf{H}' \cdot d\mathbf{l} = 4\pi\alpha\beta \int_s \mathbf{j}' \cdot d\mathbf{s}, \qquad (3)$$

$$\oint \mathbf{E}' \cdot d\mathbf{l} = -\alpha\delta \int_s \frac{\partial \mathbf{B}'}{\partial t} \cdot d\mathbf{s}, \qquad (4)$$

$$D' = \frac{E'}{\alpha\gamma c^2} + 4\pi \frac{\alpha}{\gamma} P' = \frac{\kappa}{\alpha\gamma c^2} E', \qquad (5)$$

$$B' = \frac{H'}{\beta\delta} + 4\pi \frac{\beta}{\delta} I' = \frac{\mu}{\beta\delta} H'. \qquad (6)$$

The e.s.u. system of equations is obtained by making $\alpha = \gamma = 1/c$, $\beta = \delta = c$, giving

$$E_s = \frac{q_s}{r^2}, \qquad (1s) \qquad\qquad H_s = c^2 \frac{m_s}{r^2}, \qquad (2s)$$

$$\oint \mathbf{H}_s \cdot d\mathbf{l} = 4\pi \int_s \mathbf{j}_s \cdot d\mathbf{s}, \qquad (3s)$$

$$\oint \mathbf{E}_s \cdot d\mathbf{l} = -\int_s \frac{\partial \mathbf{B}_s}{\partial t} \cdot d\mathbf{s}, \qquad (4s)$$

$$D_s = E_s + 4\pi P_s = \kappa E_s, \qquad (5s)$$

$$B_s = \frac{H_s}{c^2} + 4\pi I_s = \frac{\mu}{c^2} H_s. \qquad (6s)$$

The fundamental equations are symmetrical neither in e.m.u. nor in e.s.u. A much superior set of c.g.s. units has been proposed by Heaviside and Lorentz. These are known as the *Heaviside-*

Lorentz units and are designated by the abbreviation h.l.u. They have the advantage of making apparent the symmetry existing between electrical quantities on the one hand and magnetic quantities on the other by putting corresponding equations in the same form, as well as of getting rid of a number of numerical and dimensional constants appearing in the e.s.u. and e.m.u. systems. The h.l.u. system is used almost universally by writers on electromagnetic theory at the present time.

We shall designate quantities measured in h.l.u. by the subscript l. In this system we put

$$\alpha = \frac{1}{c\sqrt{4\pi}}, \qquad \beta = \frac{1}{\sqrt{4\pi}}, \qquad \gamma = \frac{\sqrt{4\pi}}{c}, \qquad \delta = \sqrt{4\pi},$$

getting

$$E_l = \frac{q_l}{4\pi r^2}, \qquad (1l) \qquad\qquad H_l = \frac{m_l}{4\pi r^2}, \qquad (2l)$$

$$\oint \mathbf{H}_l \cdot d\mathbf{l} = \frac{1}{c}\int_s \mathbf{j}_l \cdot d\mathbf{s}, \qquad (3l)$$

$$\oint \mathbf{E}_l \cdot d\mathbf{l} = -\frac{1}{c}\int_s \frac{\partial \mathbf{B}_l}{\partial t} \cdot d\mathbf{s}, \qquad (4l)$$

$$D_l = E_l + P_l = \kappa E_l, \qquad (5l)$$

$$B_l = H_l + I_l = \mu H_l. \qquad (6l)$$

It is to be noticed that the force between two charges measured in h.l.u. is

$$F = \frac{q_l q_l'}{4\pi r^2}. \qquad (108\text{–}2)$$

Therefore the Heaviside-Lorentz unit of charge is that charge which repels an equal like charge at a distance of 1 cm *in vacuo* with a force of $1/4\pi$ dyne. Similar remarks apply to the unit of pole strength.

Gauss' law (14–4) takes the form

$$\int_s \mathbf{D}_l \cdot d\mathbf{s} = \int_\tau \rho_l d\tau \qquad (108\text{–}3)$$

in h.l.u., so while we have introduced a factor $1/4\pi$ in Coulomb's law we have rid ourselves of 4π in Gauss' law. As Gauss' law is a more useful form of the law of force than Coulomb's, this constitutes a distinct gain in simplicity.

The following table gives conversion relations between the more important electromagnetic quantities in h.l.u., e.m.u. and e.s.u.

Charge,	$q_l = c\sqrt{4\pi}q_m = \sqrt{4\pi}q_s,$
Current,	$i_l = c\sqrt{4\pi}i_m = \sqrt{4\pi}i_s,$
Electric Intensity (e.m.f. and elect. pot.),	$E_l = \dfrac{1}{c\sqrt{4\pi}}E_m = \dfrac{1}{\sqrt{4\pi}}E_s,$
Polarization,	$P_l = c\sqrt{4\pi}P_m = \sqrt{4\pi}P_s,$
Electric Displacement,	$D_l = \dfrac{c}{\sqrt{4\pi}}D_m = \dfrac{1}{\sqrt{4\pi}}D_s,$
Magnetic Intensity (m.m.f. and mag. pot.),	$H_l = \dfrac{1}{\sqrt{4\pi}}H_m = \dfrac{1}{c\sqrt{4\pi}}H_s,$
Intensity of Magnetization,	$I_l = \sqrt{4\pi}I_m = c\sqrt{4\pi}I_s,$
Magnetic Induction,	$B_l = \dfrac{1}{\sqrt{4\pi}}B_m = \dfrac{c}{\sqrt{4\pi}}B_s,$
Resistance,	$R_l = \dfrac{1}{4\pi c^2}R_m = \dfrac{1}{4\pi}R_s,$
Capacity,	$C_l = 4\pi c^2 C_m = 4\pi C_s,$
Inductance,	$L_l = \dfrac{1}{4\pi c^2}L_m = \dfrac{1}{4\pi}L_s.$

The practical units are not a c.g.s. system but are obtained from the e.m.u. system by taking as unit of length one earth quadrant [$(10)^9$ cm], as unit of mass the eleventh-gram [$(10)^{-11}$ gm], and as unit of time the second. Therefore the same set of fundamental equations (1m)–(5m) hold for practical units as for e.m.u.

Certain equations, which are really definitions of one of the quantities involved, hold in all systems of units, including prac-

tical units. In addition to (108–1) the following are among these:

$$E = -\frac{\partial V}{\partial l}, \qquad C = \frac{q}{V}, \qquad i = \int_s \mathbf{j}\cdot d\mathbf{s},$$

$$\mathcal{E} = \oint \mathbf{E}\cdot d\mathbf{l}, \qquad R = \frac{\mathcal{E}}{i}, \qquad N = \int_s \mathbf{B}\cdot d\mathbf{s},$$

$$L = -\frac{\mathcal{E}}{\frac{di}{dt}}.$$

Finally, κ and μ, being defined as the ratios of D_s to E_s and of B_m to H_m, respectively, are pure numbers, and therefore the same in all systems of units. On page xii is given a conversion table between quantities measured in e.m.u., e.s.u. and practical units.

Problem 108a. Verify the table on page 442.

Problem 108b. Draw up a table showing the relations between quantities as measured in h.l.u. and as measured in practical units.

CHAPTER XIII

ALTERNATING CURRENTS

109. Simple A. C. Circuits. — When a periodic or alternating electromotive force is introduced into a circuit an alternating current results. The e.m.f. need not be sinusoidal in form although it very frequently is so in practice. We may restrict our analysis to the sinusoidal case, however, for a non-sinusoidal periodic e.m.f. may be expressed as an infinite series of sinusoidal terms whose frequencies are integral multiples of the fundamental frequency of the e.m.f. The current is then the sum of the currents corresponding to the separate terms in the infinite series. An infinite series of the sort just described is called a *Fourier series*.

We have already made a preliminary study of a simple series circuit with resistance, inductance and capacity in article 90. Denoting these quantities by R, L and C, respectively, an impressed electromotive force $\mathcal{E}_0 \sin \omega t$ gives rise to a current

$$i = i_0 \sin (\omega t - \phi),$$

where

$$i_0 \equiv \frac{\mathcal{E}_0}{\sqrt{R^2 + \left(L\omega - \frac{1}{C\omega}\right)^2}}, \qquad \tan \phi \equiv \frac{L\omega - \frac{1}{C\omega}}{R}.$$

These equations were deduced on the assumption that the circuit elements consist of one unit each, but, as was pointed out in article 89, R, L and C may also represent resultant values if each element is a combination of units. We know how to calculate these values for series and parallel arrangements of resistances (art. 52) and of capacities (art. 18).

When inductances are placed in series the total induced e.m.f. is the sum of the separate e.m.f.'s. Thus, if i is the current

flowing through the inductances and L is the resultant or equivalent inductance,

$$L\frac{di}{dt} = L_1\frac{di}{dt} + L_2\frac{di}{dt} + \cdots L_k\frac{di}{dt} = \left[\sum_1^k L_j\right]\frac{di}{dt},$$

and therefore

$$L = \sum_1^k L_j. \qquad (109\text{–}1)$$

With inductances in parallel the induced e.m.f. is the same for each inductance, while the currents i_1, i_2, $\cdots$ i_k are in general different. Then

$$L\frac{di}{dt} = L_1\frac{di_1}{dt} = L_2\frac{di_2}{dt} = \cdots L_k\frac{di_k}{dt}, \qquad (109\text{–}2)$$

where $i = i_1 + i_2 + \cdots i_k$ is the total current. Writing (109–2) in the form

$$\frac{L}{L_1}\frac{di}{dt} = \frac{di_1}{dt}, \quad \frac{L}{L_2}\frac{di}{dt} = \frac{di_2}{dt}, \quad \cdots \frac{L}{L_k}\frac{di}{dt} = \frac{di_k}{dt};$$

and adding, gives

$$\left(\frac{L}{L_1} + \frac{L}{L_2} + \cdots \frac{L}{L_k}\right)\frac{di}{dt} = \frac{di_1}{dt} + \frac{di_2}{dt} + \cdots \frac{di_k}{dt}.$$

As the right-hand member of this equation is equal to $\frac{di}{dt}$, we must have

$$\frac{L}{L_1} + \frac{L}{L_2} + \cdots \frac{L}{L_k} = 1,$$

or, finally,

$$\frac{1}{L} = \sum_1^k \frac{1}{L_j}. \qquad (109\text{–}3)$$

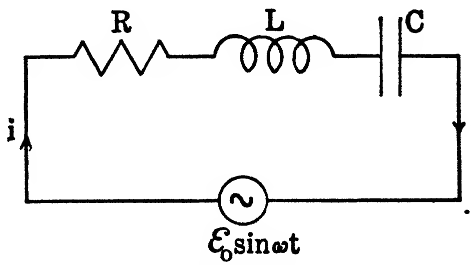


Fig. 230

Thus self-inductances combine, when arranged in series or in parallel, in the same way as resistances.

Let us now return to the original circuit (Fig. 230) and consider the relation between current and e.m.f. for each element.

The e.m.f. across the resistance is given by

$$\mathcal{E}_R = Ri = Ri_0 \sin(\omega t - \phi), \tag{109–4}$$

which is simply Ohm's law.

The e.m.f. measured externally across the inductance is, as shown in article 87,

$$\mathcal{E}_L = L\frac{di}{dt} = L\omega i_0 \cos(\omega t - \phi). \tag{109–5}$$

The amplitude of this e.m.f. is therefore obtained by multiplying the amplitude of the current by the magnitude of the inductive reactance. This is analogous to Ohm's law for the resistance. In this case, however, there is a difference in phase between e.m.f. and current. In fact, since $\cos(\omega t - \phi) = \sin(\omega t - \phi + \pi/2)$, i lags $\pi/2$ radians behind $\mathcal{E}_L$. This means that the current in the inductance reaches its maximum value a quarter of a period after the e.m.f. across the inductance.

The e.m.f. across the capacity is

$$\mathcal{E}_C = \frac{q}{C} = -\frac{1}{C\omega} i_0 \cos(\omega t - \phi), \tag{109–6}$$

since $q = \int i dt$. The amplitude of this e.m.f. is equal to the amplitude of the current multiplied by the magnitude of the capacitative reactance. As regards phase, i is evidently $\pi/2$ radians ahead of $\mathcal{E}_C$. The phase relations in (109–5) and (109–6) are often expressed by saying that an inductance draws a lagging current and a capacity draws a leading current.

For the whole circuit, the amplitude of the e.m.f. is

$$\mathcal{E}_0 = Zi_0, \qquad Z \equiv \sqrt{R^2 + \left(L\omega - \frac{1}{C\omega}\right)^2}, \tag{109–7}$$

from (90–3) and (90–4); and, denoting the reactance by X, the phase difference is given by

$$\tan\phi = \frac{X}{R}, \qquad X \equiv L\omega - \frac{1}{C\omega}. \tag{109–8}$$

A positive value of ϕ indicates that the current lags behind the e.m.f.

Evidently (109–7) and (109–8) include the previous results. For let us reduce the circuit to a resistance only by setting $L = 0$, $C = \infty$. Then X and ϕ become zero, so that the amplitude of the e.m.f. across R is Ri_0 and the phase difference is zero. Similarly with $R = 0$, $C = \infty$, the amplitude of the e.m.f. is $L\omega i_0$ and the phase difference is $\pi/2$, while with $R = 0$, $L = 0$, the values are $(1/C\omega)i_0$ and $-\pi/2$.

We may also find the relation between current and e.m.f. for any pair of circuit elements. For example, putting $C = \infty$, we see that the amplitude of the e.m.f. across a resistance and an inductance in series is $\sqrt{R^2 + L^2\omega^2}i_0$ and the angle of lag is arc tan $(L\omega/R)$.

All the foregoing relations between current and e.m.f. are valid for effective or r.m.s. values, since to obtain these it is only necessary to multiply the equations through by $1/\sqrt{2}$.

Problem 109a. Inductances of 2 millihenries and 5 millihenries respectively are connected in parallel. The combination is placed in series with a resistance of 10 ohms and a 1000-cycle e.m.f. of 100 volts. Find the amplitude and angle of lag of the current in the resistance and in each inductance. Ans. 7.44 amp, 5.31 amp, 2.13 amp; 0.731 radian.

Problem 109b. Find the amplitude and phase angle (relative to the impressed e.m.f.) of the e.m.f.'s across the resistance and the inductances in the preceding problem. Ans. 74.4 volt, $-$ 0.731 radian; 66.8 volt, 0.839 radian.

110. Divided Circuits. — Often circuits are divided into several branches. A simple illustration of a divided circuit is given in Fig. 231. To find the total current i in this circuit, denote the current through L by i_L and that through C by i_C. Also let the charge on C be q_C. Since the circuit is divided there are two circuit equations, corresponding to (87–7) and (88–4), respectively. That is,

$$L\frac{di_L}{dt} + Ri = \mathcal{E}_0 \sin \omega t,$$

$$Ri + \frac{q_C}{C} = \mathcal{E}_0 \sin \omega t.$$

Differentiating the second equation twice, multiplying through by LC and adding to the first equation, we obtain

$$RLC\frac{d^2i}{dt^2} + L\frac{di}{dt} + Ri = \mathcal{E}_0(1 - LC\omega^2)\sin\omega t, \qquad (110\text{–}1)$$

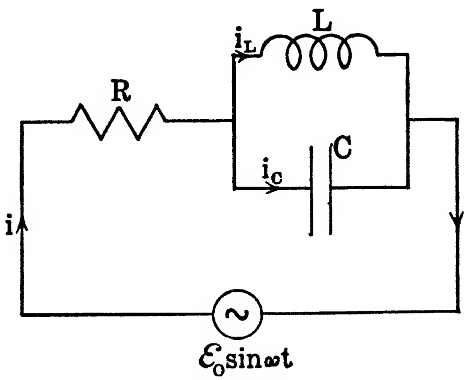


FIG. 231

since $i_L + i_C = i$. As in the case of (90–1) we know the solution of (110–1) has the form,

$$i = i_0 \sin(\omega t - \phi).$$

Substituting in the differential equation and arranging terms we have

$$\{(-RLC\omega^2\cos\phi + L\omega\sin\phi + R\cos\phi)i_0 - \mathcal{E}_0(1 - LC\omega^2)\}\sin\omega t$$
$$+\{RLC\omega^2\sin\phi + L\omega\cos\phi - R\sin\phi\}i_0\cos\omega t = 0,$$

showing that to satisfy (110–1) we must have

$$i_0 = \frac{\mathcal{E}_0}{R\cos\phi + \left(\frac{L\omega}{1 - LC\omega^2}\right)\sin\phi}, \quad \tan\phi = \frac{1}{R}\left(\frac{L\omega}{1 - LC\omega^2}\right).$$

The values of $\sin\phi$ and $\cos\phi$ may be obtained from the expression for $\tan\phi$, and we find that

$$i_0 = \frac{\mathcal{E}_0}{\sqrt{R^2 + \left(\frac{L\omega}{1 - LC\omega^2}\right)^2}}, \qquad (110\text{–}2)$$

and hence

$$Z = \sqrt{R^2 + \left(\frac{L\omega}{1 - LC\omega^2}\right)^2}. \qquad (110\text{–}3)$$

Inspection of (110–3) shows that $Z = \infty$ for a frequency $\nu_r = 1/2\pi\sqrt{LC}$. This is the exact antithesis of the series circuit case where the impedance has a minimum value for this frequency. Thus the frequency which makes the series circuit resonant makes the divided circuit *antiresonant*. Antiresonant circuits are often useful when it is necessary to eliminate certain definite frequencies in a current supply.

Usually divided circuits are more complicated than the one just discussed. For example, the resistance R_L of an inductance coil often cannot be neglected. Consider the circuit shown in Fig. 232. This may be solved by the same method as before,

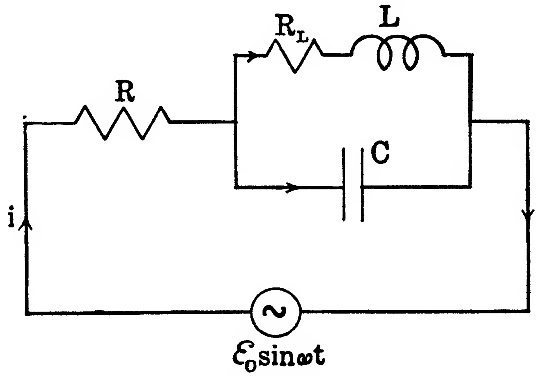


FIG. 232

but, as the reader may easily convince himself, the analysis is quite involved. We therefore defer the solution of this and more elaborate circuits and turn our attention to the development of a new and more powerful method of attacking circuit problems which depends on the use of complex quantities.

111. Complex Quantities. — Although the square roots of negative numbers have no existence in the domain of real quantities, we often find it necessary to deal with them in analytical

problems. A quantity of this type, called an *imaginary*, may evidently be expressed in the form $\pm\,\mathrm{i}y$ where $\mathrm{i} \equiv \sqrt{-1}$ and y is the square root of the magnitude of the negative number. The combination of a real quantity and an imaginary is called a *complex quantity*. Thus $x + \mathrm{i}y$ and its *conjugate* $x - \mathrm{i}y$ are complex quantities. We shall use **black face** type, as in the case of vectors, to distinguish a complex quantity. For example,

$$\mathbf{z} = x + \mathrm{i}y. \tag{111–1}$$

Complex numbers obey all the ordinary laws of algebraic manipulation. Thus to add, and to subtract,

$$\left.\begin{aligned} \mathbf{z} + \mathbf{z}' &= (x + \mathrm{i}y) + (x' + \mathrm{i}y') = (x + x') + \mathrm{i}(y + v'), \\ \mathbf{z} - \mathbf{z}' &= (x + \mathrm{i}y) - (x' + \mathrm{i}y') = (x - x') + \mathrm{i}(y - y'). \end{aligned}\right\} \tag{111–2}$$

To multiply, and to divide,

$$\left.\begin{aligned} \mathbf{z}\mathbf{z}' &= (x + \mathrm{i}y)\,(x' + \mathrm{i}y') = (xx' - yy') + \mathrm{i}(xy' + x'y), \\ \frac{\mathbf{z}}{\mathbf{z}'} &= \frac{x + \mathrm{i}y}{x' + \mathrm{i}y'} = \frac{(x + \mathrm{i}y)(x' - \mathrm{i}y')}{x'^2 + y'^2} \\ &= \left(\frac{xx' + yy'}{x'^2 + y'^2}\right) - \mathrm{i}\left(\frac{xy' - x'y}{x'^2 + y'^2}\right). \end{aligned}\right\} \tag{111–3}$$

It is important to observe that since the real and the imaginary parts of a complex equation are entirely independent, such an equation represents two real equations. If we set

$$\mathbf{z} + \mathbf{z}' = \mathbf{z}\mathbf{z}',$$

for instance, we must have

$$x + x' = xx' - yy',$$
$$y + y' = xy' + x'y.$$

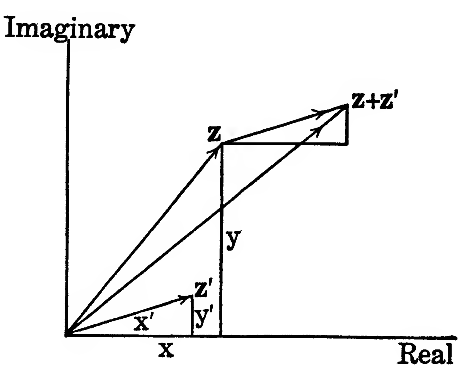


FIG. 233

We may represent $\mathbf{z}$ graphically by a point in the XY plane, taking the X axis as the *real axis* and the Y axis as the *imaginary axis*. In Fig. 233 are shown $\mathbf{z}$, $\mathbf{z}'$ and $\mathbf{z} + \mathbf{z}'$. Evidently, if we

think of the lines from the origin to the given points rather than of the points themselves as representing the complex quantities, the addition of complex quantities is, graphically, exactly like the addition of vectors. That is, the directed lines are placed head to tail and a closing line drawn to obtain the sum. Similarly, subtraction is effected graphically by reversing the direction of the representative line and then adding.

The vector properties of complex quantities suggest that it may be convenient to express the latter in terms of amplitude and direction as well as in terms of real and imaginary parts. Let us use polar coordinates z, ϕ. Then

$$x = z \cos \phi, \qquad y = z \sin \phi; \tag{111-4}$$

and, conversely,

$$z = \sqrt{x^2 + y^2}, \qquad \tan \phi = \frac{y}{x}. \tag{111-5}$$

These relations are illustrated in Fig. 234, which shows the same quantities as Fig. 233, in terms of z and ϕ.

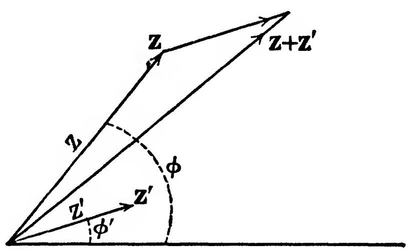


FIG. 234

The polar form $\mathbf{z} = z(\cos \phi + \mathrm{i} \sin \phi)$ leads us to an important result, for

$$\cos \phi + \mathrm{i} \sin \phi = e^{\mathrm{i}\phi}, \tag{111-6}$$

as may be seen by expanding all terms in infinite series, e being $2.718 \cdots$, the base of the Naperian logarithms, as usual. Therefore

$$\mathbf{z} = z e^{\mathrm{i}\phi}. \tag{111-7}$$

The exponential form of complex quantities is very convenient for multiplication and division, just as the rectangular form is especially adapted to addition and subtraction. Thus,

$$\left.\begin{aligned} \mathbf{z}\mathbf{z}' &= (z e^{\mathrm{i}\phi})(z' e^{\mathrm{i}\phi'}) = z z' e^{\mathrm{i}(\phi+\phi')}, \\ \frac{\mathbf{z}}{\mathbf{z}'} &= \frac{z e^{\mathrm{i}\phi}}{z' e^{\mathrm{i}\phi'}} = \frac{z}{z'} e^{\mathrm{i}(\phi-\phi')}. \end{aligned}\right\} \tag{111-8}$$

Graphically, multiplication or division by a complex number is represented by a change of amplitude and a rotation of the quantity affected. For example, multiplying $\mathbf{z}'$ by $\mathbf{z}$ increases the amplitude of the former by a factor z and rotates it through an angle ϕ. A special case of interest is multiplication by i. Since $\mathrm{i} = e^{\mathrm{i}(\pi/2)}$, this merely causes rotation through $\pi/2$ radians. Similarly division by i causes a rotation of $-\pi/2$ radians.

From (111–2) and (111–8) we may formulate the rules for combining complex quantities:

1. *In addition (subtraction) the components of the sum (difference) are equal respectively to the sum (difference) of the components.*

2. *In multiplication (division) the amplitude of the product (quotient) is equal to the product (quotient) of the amplitudes, and the angle of the product (quotient) is equal to the sum (difference) of the angles.*

In the course of complex analysis we express quantities in rectangular or in exponential form according to convenience, passing from one to the other by means of (111–4) and (111–5). Numerous examples of this procedure will be found in following articles.

Problem 111a. Show that

$$\sqrt{x+\mathrm{i}y} = \pm\frac{1}{\sqrt{2}}\left(\sqrt{\sqrt{x^2+y^2}+x}+\mathrm{i}\sqrt{\sqrt{x^2+y^2}-x}\right).$$

Problem 111b. Prove that if n is an integer

$$\sqrt[n]{\mathrm{i}} = \cos\left(\frac{\frac{\pi}{2}+2\pi p}{n}\right)+\mathrm{i}\sin\left(\frac{\frac{\pi}{2}+2\pi p}{n}\right); \qquad p = 0, 1, \cdots (n-1)$$

Problem 111c. Show that

$$(\cos\phi+\mathrm{i}\sin\phi)^n = \cos n\phi+\mathrm{i}\sin n\phi.$$

Problem 111d. Find the value of i^{i}. Ans. 0.208.

Problem 111e. Find the real part and also the amplitude of

$$\left(\frac{1-2\mathrm{i}}{4+2\mathrm{i}}\right)(2+5\mathrm{i}).$$

Ans. $5/2$, $\sqrt{29}/2$.

112. Application of Complex Quantities to A. C. Circuits. — In order to see how complex quantities may be used in solving a.c. circuit problems, consider again the differential equation of a simple series circuit (art. 90). This equation is

$$L\frac{d^2i}{dt^2}+R\frac{di}{dt}+\frac{1}{C}i=\frac{d\mathcal{E}}{dt}.$$

Suppose we indicate the solution corresponding to $\mathcal{E} = \mathcal{E}_0 \cos \omega t$ by i_x and that corresponding to $\mathcal{E} = \mathcal{E}_0 \sin \omega t$ by i_y. Then

$$L\frac{d^2i_x}{dt^2}+R\frac{di_x}{dt}+\frac{1}{C}i_x=\frac{d}{dt}(\mathcal{E}_0\cos\omega t),$$

$$L\frac{d^2i_y}{dt^2}+R\frac{di_y}{dt}+\frac{1}{C}i_y=\frac{d}{dt}(\mathcal{E}_0\sin\omega t).$$

Multiplying the second of these equations by i and adding it to the first gives

$$L\frac{d^2\mathbf{i}}{dt^2}+R\frac{d\mathbf{i}}{dt}+\frac{1}{C}\mathbf{i}=\frac{d}{dt}(\mathcal{E}_0e^{\mathrm{i}\omega t}),\tag{112–1}$$

where $\mathbf{i} = i_x + \mathrm{i}i_y$. Thus, in solving the circuit problem, if we replace the actual sinusoidal e.m.f. by a corresponding complex e.m.f., we obtain a complex current whose real part is the current due to an applied cosine e.m.f. and whose imaginary part is the current due to an applied sine e.m.f. The utility of this procedure lies in the fact that it is simpler to solve the differential equation in the complex form than in the real form. This is not of great importance in the case of (112–1), but the method is applicable to all a.c. circuit problems and is of very great value in the more complicated cases.

Let us now solve (112–1). The solution must be of the form $\mathbf{i}_0e^{\mathrm{i}\omega t}$, since current and e.m.f. have the same frequency. Substituting in the differential equation,

$$\left(-L\omega^2+R\mathrm{i}\omega+\frac{1}{C}\right)\mathbf{i}_0e^{\mathrm{i}\omega t}=\mathcal{E}_0\mathrm{i}\omega e^{\mathrm{i}\omega t},$$

so, in order that (112–1) may be satisfied, we must have

$$\mathbf{i}_0 = \frac{\mathcal{E}_0}{R + \mathrm{i}\left(L\omega - \dfrac{1}{C\omega}\right)} \cdot \qquad (112\text{–}2)$$

The denominator of the right-hand side of (112–2) is the *complex impedance* **Z** of the circuit. Its amplitude

$$Z = \sqrt{R^2 + \left(L\omega - \frac{1}{C\omega}\right)^2}$$

is the real impedance to which we are accustomed, and its angle ϕ, given by

$$\tan\phi = \frac{\left(L\omega - \dfrac{1}{C\omega}\right)}{R},$$

is the angle of lag.

Expressing **Z** in exponential form,

$$\mathbf{i} = \frac{\mathcal{E}_0}{Ze^{\mathrm{i}\phi}}e^{\mathrm{i}\omega t} = \frac{\mathcal{E}_0}{Z}e^{\mathrm{i}(\omega t - \phi)}. \qquad (112\text{–}3)$$

The actual current corresponding to an applied e.m.f. $\mathcal{E}_0 \sin \omega t$ is the imaginary part of (112–3), that is,

$$i = \frac{\mathcal{E}_0}{Z}\sin(\omega t - \phi),$$

which is the value previously found in (90–2). Since actual values are obtained so simply from the complex expressions, circuit problems often reduce merely to a calculation of impedance.

In complex form the relation between current and e.m.f. is

$$\boldsymbol{\mathcal{E}} = \mathbf{Z}\mathbf{i}. \qquad (112\text{–}4)$$

This expression holds not only for the simple case of (112–3) but for any circuit, or for any part of a circuit. The value of **Z** depends, of course, on the arrangement of circuit elements involved, but it is always of the form

$$\mathbf{Z} = \hat{R} + \mathrm{i}\hat{X}, \qquad (112\text{–}5)$$

where $\hat{R}$ is the equivalent series resistance, and $\hat{X}$ is the equivalent reactance. Comparing (112–4) with Ohm's law in real form, $\mathcal{E} = Ri$, it is evident that in a complex a.c. circuit impedances behave exactly as resistances do in a real d.c. circuit. Thus, impedances in series add directly, and impedances in parallel add by reciprocals. Since the impedances of the elements R, L and C are R, $\mathrm{i}L\omega$ and $-\mathrm{i}\dfrac{1}{C\omega}$, respectively, we can easily calculate the impedance of any series and parallel combination of them.

We are now able to deal with the circuit of Fig. 232. The impedance of the upper branch of the divided circuit is $R_L + \mathrm{i}L\omega$ and that of the lower $-\mathrm{i}\dfrac{1}{C\omega}$. Therefore the impedance of the divided portion of the circuit is

$$\frac{1}{\dfrac{1}{R_L + \mathrm{i}L\omega} + \dfrac{1}{-\mathrm{i}\dfrac{1}{C\omega}}} = \frac{R_L + \mathrm{i}L\omega}{(1 - LC\omega^2) + \mathrm{i}R_L C\omega}.$$

Multiplying numerator and denominator by the conjugate of the latter reduces the expression to the standard form

$$\frac{R_L}{(1 - LC\omega^2)^2 + R_L^2C^2\omega^2} + \mathrm{i}\frac{L\omega(1 - LC\omega^2) - R_L^2C\omega}{(1 - LC\omega^2)^2 + R_L^2C^2\omega^2},$$

and **Z** is obtained by adding R. Setting

$$\hat{R} = R + \frac{R_L}{(1 - LC\omega^2)^2 + R_L^2C^2\omega^2},$$

$$\hat{X} = \frac{L\omega(1 - LC\omega^2) - R_L^2C\omega}{(1 - LC\omega^2)^2 + R_L^2C^2\omega^2},$$

the amplitude and angle of **Z** are given by

$$Z = \sqrt{\hat{R}^2 + \hat{X}^2}, \qquad \tan\phi = \frac{\hat{X}}{\hat{R}}.$$

If the frequency or the circuit elements are adjusted for anti-

resonance, that is, so that $1 - LC\omega^2 = 0$, we have

$$Z = \sqrt{\left(R + \frac{1}{R_L C^2\omega^2}\right)^2 + \frac{1}{C^2\omega^2}}.$$

This is not infinite as was the case for the circuit of Fig. 231, but is large when R_L is small and converges to infinity as R_L converges to zero.

Returning to (112–4), we observe that while this relation is of convenient form when circuit elements or parts of circuits are in series, an expression of the form

$$\mathbf{i} = \mathbf{Y}\boldsymbol{\mathcal{E}}, \tag{112–6}$$

is better suited for parallel combinations. We therefore define the *complex admittance* $\mathbf{Y}$ as the reciprocal of $\mathbf{Z}$. As the admittances of the elements R, L and C are respectively $\frac{1}{R}$, $-\mathrm{i}\frac{1}{L\omega}$ and $\mathrm{i}C\omega$ the total admittance of R, L and C in parallel is

$$\mathbf{Y} = \frac{1}{R} - \mathrm{i}\left(\frac{1}{L\omega} - C\omega\right) = G - \mathrm{i}B, \tag{112–7}$$

where G is called the *conductance* and B the *susceptance*. In the exponential form

$$\mathbf{Y} = Ye^{-\mathrm{i}\phi}, \tag{112–8}$$

where

$$Y = \sqrt{G^2 + B^2}, \qquad \tan\phi = \frac{B}{G}.$$

The admittance for a more complicated combination of circuit elements always has the general form

$$\mathbf{Y} = \hat{G} - \mathrm{i}\hat{B}, \tag{112–9}$$

in which $\hat{G}$ is the equivalent parallel conductance and $\hat{B}$ is the equivalent susceptance. For example, the admittance corresponding to $\mathbf{Z} = R + \mathrm{i}X$ is

$$\mathbf{Y} = \frac{1}{R + \mathrm{i}X} = \frac{R - \mathrm{i}X}{R^2 + X^2} = \frac{R}{Z^2} - \mathrm{i}\frac{X}{Z^2}.$$

Hence, in this case,

$$\hat{G} = \frac{R}{Z^2}, \qquad \hat{B} = \frac{X}{Z^2}.$$

The relation between current and e.m.f. in any circuit may be determined by calculating either **Z** or **Y**, of course. One calculation is usually found to be simpler than the other, however.

There is one restriction on the use of complex quantities in circuit analysis which must be mentioned. They cannot be used to calculate power directly, because the real or imaginary part of $\boldsymbol{\mathcal{E}}\mathbf{i}$ is not equal to the product of the real or imaginary parts of $\boldsymbol{\mathcal{E}}$ and $\mathbf{i}$ respectively. To find the power at any instant we must find the actual e.m.f. and the actual current and multiply these real quantities together.

Problem 112a. Deduce (109–4) to (109–6), inclusive, by means of complex quantities.

Problem 112b. Find the current in the circuit of Fig. 231, using complex analysis.

Ans. $$i = \frac{\mathcal{E}_0}{\sqrt{R^2 + \left(\dfrac{L\omega}{1 - LC\omega^2}\right)^2}} \sin(\omega t - \phi),$$

$$\tan\phi \equiv \frac{1}{R}\left(\frac{L\omega}{1 - LC\omega^2}\right).$$

Problem 112c. Given a current $i = i_0 \sin(\omega t + \epsilon)$ flowing through R and L in parallel. Find the e.m.f.

Ans. $$\frac{RL\omega i_0}{\sqrt{R^2 + L^2\omega^2}} \sin(\omega t + \epsilon + \phi), \qquad \tan\phi \equiv \frac{R}{L\omega}.$$

Problem 112d. In the circuit of Fig. 232 the resistance R_L in series with L is removed, and an equal resistance R_C is placed in series with C. Find **Z**. Also, show that at antiresonance Z has the same value as when the circuit contained R_L instead of R_C. (See p. 456.)

Ans. $$\mathbf{Z} = \left[R + \frac{R_C L^2 C^2 \omega^4}{(1 - LC\omega^2)^2 + R_C^2 C^2 \omega^2}\right] + \mathrm{i}\left[\frac{L\omega(1 - LC\omega^2) + R_C^2 L C^2 \omega^3}{(1 - LC\omega^2)^2 + R_C^2 C^2 \omega^2}\right].$$

113. Graphical Representation of Electrical Quantities. — Electrical quantities in complex form may be represented

graphically, according to the method described in article 111. Graphical representations are often of great value in understanding the behavior of complicated circuits and graphical analysis is particularly useful to the engineer in designing electrical machinery and other apparatus.

Figure 235 shows an impedance

$$\mathbf{Z} = R + i\left(L\omega - \frac{1}{C\omega}\right),$$

and an admittance

$$\mathbf{Y} = \frac{1}{R} - i\left(\frac{1}{L\omega} - C\omega\right).$$

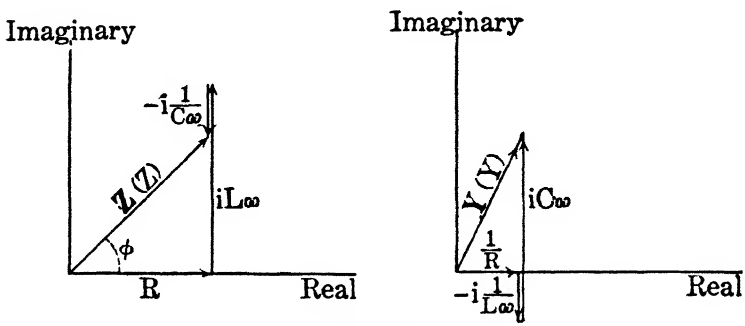


FIG. 235

A more complicated case, the impedance of the circuit of Fig. 232, is shown in Fig. 236. Line 1 is the impedance of the upper branch of the divided circuit, and its reciprocal (line 2) is the corresponding admittance. Adding $iC\omega$ we obtain line 3, the admittance of the divided part of the circuit. The reciprocal of this (line 4) is the corresponding impedance, to which we add R and obtain $\mathbf{Z}$. The diagram as drawn indicates a frequency condition below antiresonance, that is, $(1 - LC\omega^2) > 0$, and

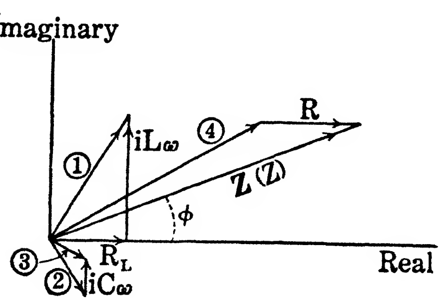


FIG. 236

shows $\phi > 0$. This agrees with the analytical result obtained on page 455 for the case where R_L is small.

Often it is more instructive to show currents and e.m.f.'s directly on the diagram, rather than impedances and admittances. In this case we usually dispense with the axes as we are interested only in the relative phases of the electrical quantities. A graphical representation of currents and e.m.f.'s is called a *vector diagram*, although the quantities represented are not true vectors in the physical sense.

To illustrate such a diagram we may use again the divided circuit to which Fig. 236 refers. Denoting the current through the upper branch by $\mathbf{i}_L$ and that through the lower branch by $\mathbf{i}_C$, the current and e.m.f. relations are represented by Fig. 237. Angles measured counterclockwise indicate a lead in phase. The diagram shows that the applied e.m.f. $\mathcal{E}$ is the sum of $R\mathbf{i}$ and the e.m.f. across the divided portion of the circuit. This last e.m.f. is, in turn, the sum of $R_L\mathbf{i}_L$ and the e.m.f. across L. As regards the currents, the total current $\mathbf{i}$, which is in phase with $R\mathbf{i}$, is the sum of $\mathbf{i}_L$, which is in phase with $R_L\mathbf{i}_L$, and $\mathbf{i}_C$, which leads the condenser e.m.f. by $\pi/2$ radians.

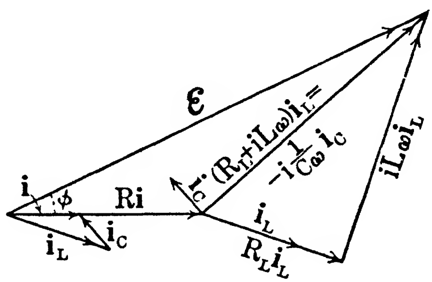


FIG. 237

Note that if we are interested in the actual or *instantaneous* values of the electrical quantities, we may draw the real and the imaginary axes and project the entire diagram on them. The projection on the real axis at any instant gives the instantaneous values for an applied cosine e.m.f., and the projection on the imaginary axis gives the corresponding values for an applied sine e.m.f. Since the phase of each of the electrical quantities contains ωt, the angles of the vectors increase uniformly with time, which is equivalent to a rotation of the vector diagram about the origin. The variation of the projections due to the rotation

represents the actual sinusoidal variation of the instantaneous values.

Problem 113a. Draw the vector diagram of a divided circuit with a resistance R_C in series with the capacity instead of R_L in series with the inductance (See problem 112*d*).

114. A.C. Networks and Kirchhoff's Laws. — When an a.c. circuit is so arranged that it cannot be analyzed into series and parallel groups of circuit elements, it constitutes an a.c. network, analogous to the d.c. network described in article 52. Evidently, as in the d.c. case, there are the two basic facts, (1) that charge cannot accumulate at a branch point in the network and (2) that the sum of the potential drops across the circuit elements around any closed path equals the sum of the applied e.m.f.'s in that path. Therefore, we may generalize Kirchhoff's laws to apply to a.c. networks. If we use real currents and e.m.f.'s the resulting equations are very awkward because of the various phase differences involved, but with complex currents and e.m.f.'s the equations are relatively simple, so we state the generalized laws as follows.

Law 1. The algebraic sum of all the complex currents meeting at a point is zero.

Law 2. The algebraic sum of the $\mathbf{Zi}$ *terms around any closed path equals the algebraic sum of the complex applied e.m.f.'s in the given path.*

Consider, for example, a network of the Wheatstone bridge type (Fig. 238), in which each branch, except the one with the e.m.f., contains an impedance of any form. As the equations arising from Kirchhoff's laws are of exactly the same form as for the d.c. bridge (art. 53), the solutions also have the same form. Hence, referring to (53–1), we see that the complex current in the gavanometer arm is

$$\mathbf{i}_g = \frac{(\mathbf{Z}_2\mathbf{Z}_3 - \mathbf{Z}_4\mathbf{Z}_1)\boldsymbol{\mathcal{E}}}{\Delta + \mathbf{Z}_g(\mathbf{Z}_1 + \mathbf{Z}_2)(\mathbf{Z}_3 + \mathbf{Z}_4)}, \qquad (114\text{–}1)$$

where

$$\Delta = \mathbf{Z}_1\mathbf{Z}_2\mathbf{Z}_3 + \mathbf{Z}_2\mathbf{Z}_3\mathbf{Z}_4 + \mathbf{Z}_3\mathbf{Z}_4\mathbf{Z}_1 + \mathbf{Z}_4\mathbf{Z}_1\mathbf{Z}_2.$$

The actual current i_g is found by taking the real or the imaginary part of (114–1), as usual.

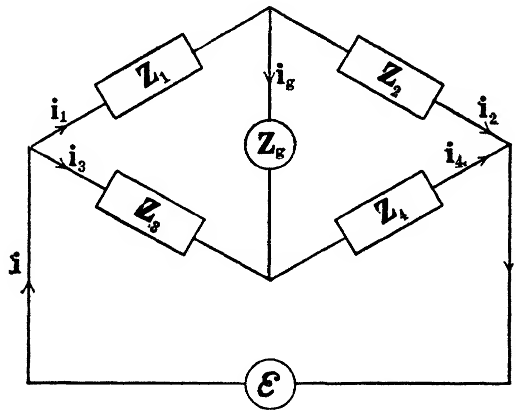


FIG. 238

When the bridge is balanced $\mathbf{i}_g = 0$, so that we must have

$$\frac{\mathbf{Z}_1}{\mathbf{Z}_2} = \frac{\mathbf{Z}_3}{\mathbf{Z}_4}. \tag{114–2}$$

This result may also be obtained by observing that, at balance,

$$\mathbf{Z}_1\mathbf{i}_1 = \mathbf{Z}_3\mathbf{i}_3, \qquad \mathbf{Z}_2\mathbf{i}_2 = \mathbf{Z}_4\mathbf{i}_4.$$

Dividing the first equation by the second gives (114–2), since $\mathbf{i}_1 = \mathbf{i}_2$, and $\mathbf{i}_3 = \mathbf{i}_4$.

A common variety of this bridge, the *inductance bridge*, is shown in Fig. 239. Usually L_1 is an unknown inductance of resistance R_1, while L_3 is a known adjustable inductance of resistance R_3. The resistances R_2 and R_4 are also known and adjustable. Equation (114–2), the condition for balance, reduces in this case to

$$\frac{R_1 + \mathrm{i}L_1\omega}{R_2} = \frac{R_3 + \mathrm{i}L_3\omega}{R_4}.$$

Equating real parts and imaginary parts, we find

$$\frac{R_1}{R_2} = \frac{R_3}{R_4}, \qquad \frac{L_1}{R_2} = \frac{L_3}{R_4}. \tag{114–3}$$

The double condition for balance is characteristic of a.c. bridges. The first relation corresponds to an ordinary d.c. balance, and is established with the aid of a constant e.m.f. and a d.c. current indicator, R_2 or R_4 being adjusted. Then

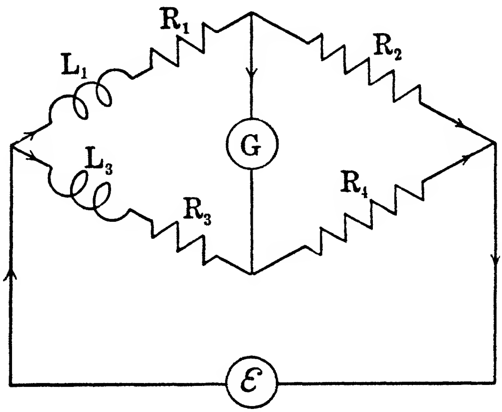


Fig. 239

an alternating e.m.f. and an a.c. current indicator are put in place of the others and the second relation in (114–3) is established by adjusting L_3, care being taken not to disturb the d.c. balance. The unknown L_1 is now determined in terms of L_3, R_2 and R_4.

Other applications of Kirchhoff's laws will be found in the next chapter, where a number of a.c. bridges are described.

115. A.C. Measuring Instruments. — Alternating currents and e.m.f.'s cannot be measured with an ordinary instrument of the D'Arsonval type because even at low frequencies the inertia of the moving element is too great to permit the latter to follow the alternating torque, and the mean torque, to which any steady deflection must be due, is zero. We may, however, use the *dynamometer* type of instrument, in which both the fixed and the moving element consist of coils. The torque at any instant on the moving part is proportional to the product of the currents in the two elements from (93–6). If these currents are drawn

from the same source the mean torque is, in general, not zero, so that a steady deflection results.

When the elements are connected in series, so that the same current i flows in each, the deflection is proportional to the mean value of i^2, which by definition (art. 90) is i_e^2, the square of the effective current. Alternating current ammeters and voltmeters are often instruments of this sort. The ammeters are made with as low resistance as possible, so they may be included in a circuit without disturbing it appreciably. The voltmeters, on the other hand, are made with high resistance, so they may be placed across any part of a circuit with negligible effect. Both ammeters and voltmeters are calibrated to indicate effective (r.m.s.) values. Because the scale is of the "squared" type small values are determined with less accuracy than large, which is often a serious defect.

In addition to the measurement of current and e.m.f., it is important to be able to determine the mean power supplied to a circuit or to part of a circuit. This can be done by means of a dynamometer instrument with one low resistance element and one high resistance element. The former is connected in series with the load and the latter across it (Fig. 240). The instantaneous torque on the moving element is therefore proportional to $\mathcal{E}i$, where $\mathcal{E}$ is the e.m.f. across the load and i is the current through it. The mean torque, and hence the deflection, is proportional to the mean value of $\mathcal{E}i$, which is the mean power.

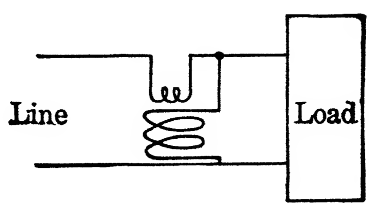


FIG. 240

The instrument just described is called a *wattmeter*. It is accurate and can be constructed for almost any power range desired, but, in common with all dynamometer type instruments, its use is limited to relatively low frequencies, not over a few hundred cycles.

There are other types of a.c. instrument beside the dynamometer. ·One of the most widely used is the *moving iron* type in

which two parallel pieces of soft iron, one fixed and one movable, are surrounded by a single coil coaxial with the moving element. Current in the coil produces a magnetic field which magnetizes the two pieces of iron. The resulting mutual repulsion causes the movable iron to be deflected. Moving iron instruments, like dynamometers, are restricted to low frequencies. For high frequencies thermocouple instruments, already described in article 64, are commonly employed.

For null measurements, such as are involved in balancing a.c. bridges, an instrument sensitive to small currents is required. In the range of audio frequencies, that is, a few hundred to a few thousand cycles, a telephone receiver serves very well. For low frequencies the best instrument is the *vibration galvanometer.* This is similar to a D'Arsonval galvanometer in that it contains a fixed magnet and a moving coil. The method of suspending the coil, however, is different. The coil, which is made very small, is held between two metal strips whose length and tension may

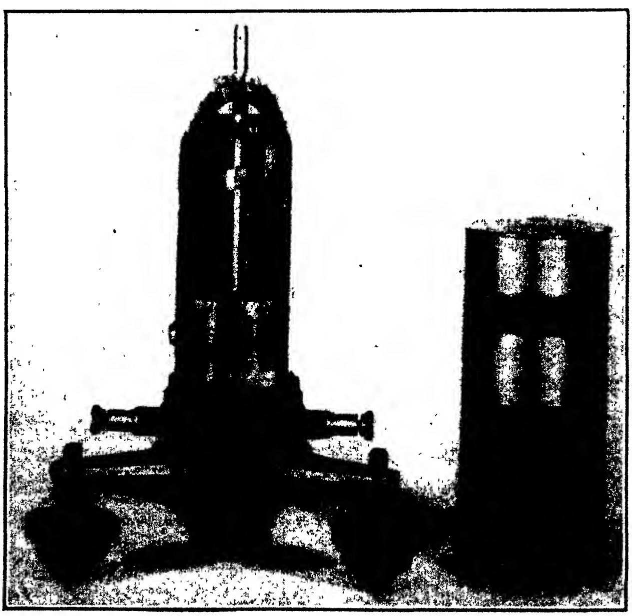


FIG. 241

be varied until the period of mechanical oscillation of the moving element is exactly equal to the period of the current in the coil. In this condition of resonance the periodic force due to a very small current builds up vibrations of considerable amplitude. A small mirror mounted on the coil reflects a spot of light on a scale. During vibration the spot of light is drawn out into a band whose length indicates the amplitude. A vibration galvanometer with the case removed is shown in Fig. 241.

For a more detailed description of a.c. instruments the reader should consult a work devoted to them. (See, for instance, F. A. Laws: *Electrical Measurements*.)

There is one important instrument, the *oscillograph*, which is of an entirely different nature from the instruments discussed above. The oscillograph is a device for examining the wave form of the current or the e.m.f. in a circuit. The type most used, due to Duddell, consists essentially of a fine strip of phosphor-bronze or similar material doubled over a pulley P (Fig. 242) and placed between the poles of a permanent magnet. A tiny mirror M is cemented across the loop formed by the metal strip.

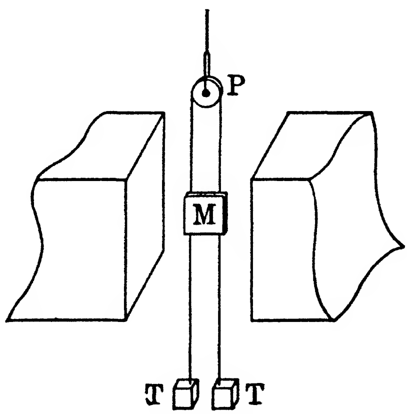


FIG. 242

Current from the source whose wave form is to be investigated is introduced into the loop through the terminal blocks T, T. As the current passes up on one side of the loop and down on the other, the forces on the two sides due to the field of the magnet are in opposite directions, causing a rotation of the mirror. By placing tension on the pulley sufficient to reduce the natural period of vibration of the mirror to one or two ten-thousandths of a second and then immersing the entire unit in oil to damp free oscillations, the motion of the mirror can be made to follow the

variation of current up to frequencies of several thousand cycles.

The deflection of the mirror is indicated by a spot of light reflected by the mirror on a screen. In order actually to depict the wave form, that is, to show deflection as a function of time, it is necessary to give the spot of light a motion perpendicular to the deflection. This is usually accomplished by reflecting the light from an auxiliary mirror or prism which rotates about an axis perpendicular to the axis of rotation of the oscillograph mirror. Ordinarily the curve representing the wave form is traced on the screen in a small fraction of a second, too fast for the eye to follow, but by driving the auxiliary mirror at the proper synchronous speed the curve is repeated rapidly so that a complete and apparently steady picture of the wave form appears.

If the wave form is to be analyzed into sinusoidal components (p. 444) it is desirable to have a permanent record of it. In this case the screen and the auxiliary mirror or prism are replaced by a film mounted on a rotating drum. The photograph so obtained is called an *oscillogram*. Figure 187 is a reproduction of an oscillogram.

116. A.C. Machinery. — Alternating current machinery may be divided into two classes, *generators* or *alternators* which convert mechanical energy into electrical, and *motors* which convert electrical energy into mechanical.

Generators. — The simplest a.c. generator consists of a coil of wire rotating in a constant magnetic field, as described in article 86. In order to obtain sufficient induced e.m.f. several coils in series, wound on an iron frame, and an equal number of field magnets are usually employed. The former comprise the *armature* and the latter the *field.* Since the armature winding is more complicated and requires better insulation than the field, it is usual practice to keep it stationary and to rotate the field. In any case the stationary part is called the *stator*, and the rotating part the *rotor.*

A typical arrangement is shown in Fig. 243. The armature coils (A), each of which contains a number of turns, lie in slots

in the frame while the field windings are placed on salient pole pieces attached to the rotating element. The d.c. current necessary to *excite* the field is supplied through brushes resting on *slip rings*.

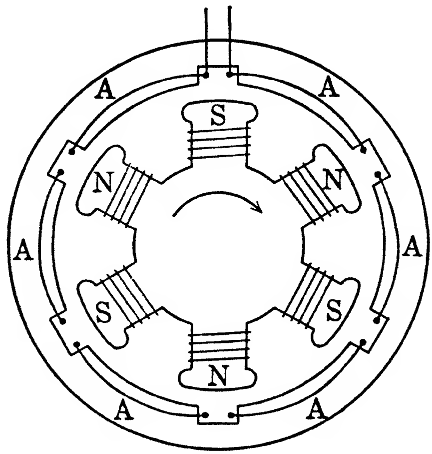


FIG. 243

The machine just illustrated is a *single-phase generator*. That is, it produces a single e.m.f. of the form $\mathcal{E} = \mathcal{E}_0 \sin \omega t$ between its terminals. For the efficient operation of motors, however, several e.m.f.'s of uniform phase difference are desirable. A *polyphase generator* is exactly like a single-phase generator except that it has as many independent sets of armature coils as there are phases, uniformly spaced around the armature frame. For example, a three-phase generator, which is the most common, produces e.m.f.'s

$$\mathcal{E}_1 = \mathcal{E}_{10} \sin \omega t,$$

$$\mathcal{E}_2 = \mathcal{E}_{20} \sin \left(\omega t + \frac{2\pi}{3} \right),$$

$$\mathcal{E}_3 = \mathcal{E}_{30} \sin \left(\omega t + \frac{4\pi}{3} \right).$$

To transmit power from a three-phase machine four wires are required, in general, one end of each phase being connected to *neutral* (Fig. 244). However, when the loads are balanced so that the currents in the three phases have the same amplitude and uniform phase difference, the neutral wire may be omitted. For the current in the neutral wire at any instant is the sum of the currents in the three phases and, under the conditions speci-

fied, this sum is zero, as shown by the vector diagram in Fig. 245. Although the neutral wire is not used for transmission it is usually brought out of the machine and connected to ground to stabilize the system.

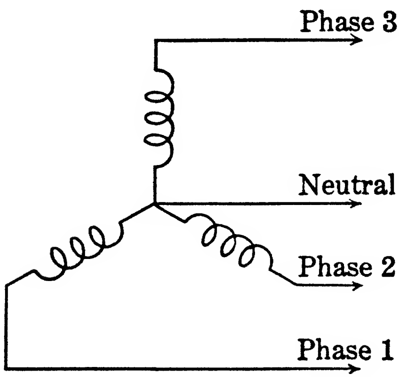


FIG. 244

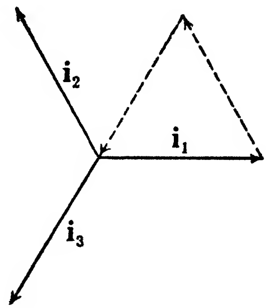


FIG. 245

Motors. — The simplest type of a.c. motor to understand is the *synchronous motor*, which is, in effect, an alternator running backwards. Consider two identical machines connected electrically phase to phase and running at the same speed, one as a generator and the other as a motor. In the generator the reaction of the stator on the rotor is such as to oppose the motion of the latter, causing the machine to absorb energy mechanically. In the motor conditions are exactly reversed, for current is everywhere in the opposite direction to that in the generator, so that the reaction of the stator on the rotor assists the motion and produces mechanical work. Evidently this state of affairs exists only when the speed of the motor is exactly that of the generator. If the load on the motor is increased until the latter is pulled out of synchronous speed, the torque falls to zero and the motor stops.

Although any alternator may be used as a synchronous motor, and *vice versa*, there are usually slight differences in the details of construction. For one thing, since the motor has no starting

torque, it must be provided with an auxiliary winding or some other device by means of which it can be brought up to speed.

The most useful type of a.c. motor is the *induction motor.* The stator of an induction motor is similar to that of a synchronous motor. The rotor, however, is entirely different, consisting merely of a core with some closed windings on it, or in its simplest form of a *squirrel cage* (Fig. 246) made by setting copper bars into the edges of two conducting disks. The rotor is thus energized by induction instead of by an auxiliary exciting current, a great practical advantage.

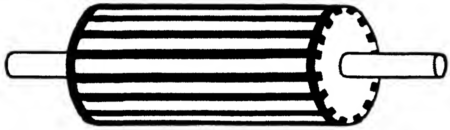


Fig. 246

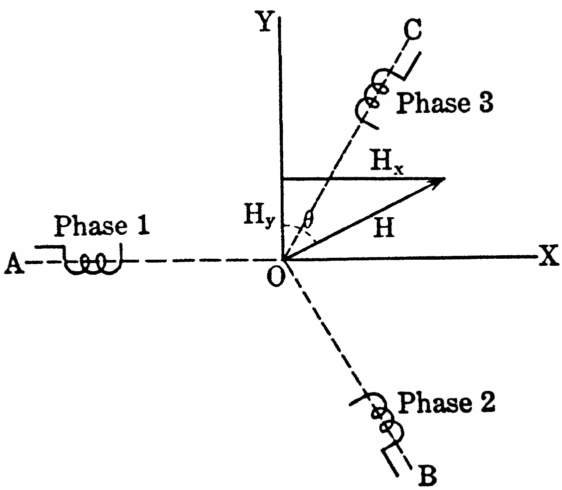


Fig. 247

To investigate the means by which rotation is produced let us take a simple case such as a three-phase machine with a single concentrated winding per phase (Fig. 247). We wish first to calculate the resultant magnetic field in the region occupied by the rotor. This is due to three fields, $H_0 \sin \omega t$ in the direction

AO, $H_0 \sin\left(\omega t + \dfrac{2\pi}{3}\right)$ in the direction BO and $H_0 \sin\left(\omega t + \dfrac{4\pi}{3}\right)$ in the direction CO.

Taking XY axes as shown in the figure, the components of the resultant field are

$$H_x = H_0\left\{\sin\omega t + \sin\left(\omega t + \frac{2\pi}{3}\right)\cos\frac{2\pi}{3} + \sin\left(\omega t + \frac{4\pi}{3}\right)\cos\frac{4\pi}{3}\right\} = \frac{3}{2}H_0 \sin\omega t,$$

$$H_y = H_0\left\{0 + \sin\left(\omega t + \frac{2\pi}{3}\right)\sin\frac{2\pi}{3} + \sin\left(\omega t + \frac{4\pi}{3}\right)\sin\frac{4\pi}{3}\right\} = \frac{3}{2}H_0 \cos\omega t.$$

Evidently then the resultant field has the constant magnitude $\frac{3}{2}H_0$. Its direction is determined by $\tan\theta = \tan\omega t$, that is, $\theta = \omega t$, which indicates a uniform rotation in a direction opposite to the progression of phases.

Now consider the effect of a rotating field on a conductor or a closed coil placed in it. Induced currents, and hence forces, are the same as for a rotating conductor in a stationary field, and by Lenz' law (art. 95) motion of a conductor in a field is always opposed. Therefore the motion of the field is opposed, which, of course, results in a torque on the conductor or coil.

Thus, in the machine described, there is a torque on the rotor in a direction opposite to the phase progression. The same thing is true in any polyphase induction motor, although when the windings for each phase are distributed around the stator the analysis of the rotating field may be very complex. The magnitude of this torque in any particular machine depends on the relative speed of field and rotor. A typical speed-torque curve is shown in Fig. 248. The torque at first increases with rotor speed, reaching a maximum at a little under synchronous speed. Thereafter the torque falls rapidly to zero as synchronism is

approached. The fact that polyphase induction motors have a starting torque constitutes one of their major advantages.

Single-phase induction motors are more difficult to analyze as the stator field is not rotating, but merely oscillatory. However, an oscillatory field may be thought of as consisting of two oppositely rotating fields of the same magnitude. There is then no resultant starting torque as the torques due to the two fields neutralize each other. If, however, the motor is started in either direction by some auxiliary device, the torque in that direction becomes greater as indicated in Fig. 248 and the motor continues to run.

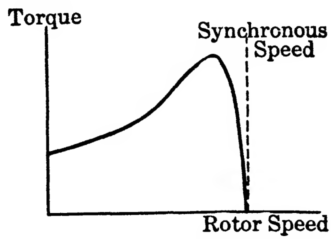


FIG. 248

Problem 116a. Show by means of a vector diagram the relation between voltage of phases to neutral and voltage between phases for a three-phase alternator. If there are 220 volts r.m.s. between phases, what is the voltage to neutral? Ans. 127.0 volt.

Problem 116b. A single-phase six-pole synchronous motor (like the alternator in Fig. 243) is driven by 60-cycle current. What is its speed of rotation? If the armature coils are reconnected for three-phase operation with a two-pole rotor at what speed will the motor run? Ans. 1200 r.p.m., 3600 r.p.m.

Problem 116c. Given a simplified induction motor consisting of a single flat coil of area A, resistance R and self-inductance L, free to revolve in a rotating magnetic field of magnitude H. Prove that the mean torque on the coil in the direction of rotation of the field is

$$\frac{H^2A^2\omega R}{2(R^2 + L^2\omega^2)},$$

where ω is the angular velocity of the field relative to the coil. Show that this leads to a speed-torque curve like that in Fig. 248.

117. Transformers. — It is often desirable to change the relative magnitudes of current and e.m.f. in a.c. systems. For instance, the generation of electrical power is most conveniently effected at a few thousand volts with large currents. On the other hand, in the transmission of power, especially over long

distances, high voltages and small currents must be used to avoid excessive ohmic losses, for the latter are of the form Ri^2 and so diminish rapidly as the current is decreased. Finally, devices using power, such as motors, are usually designed for a few hundred volts.

Transformations of current and e.m.f. of the sort indicated are readily made by means of *transformers*, which consist essentially of two coils wound on a laminated iron core, so arranged that the magnetic coupling between them is as nearly perfect as possible. Power is delivered to the *primary* winding and withdrawn from the *secondary*, being transferred from one to the other by electromagnetic induction. Figure 249 illustrates a common type of transformer construction in which the windings are superposed and the core is made in the form of a frame so designed that there is a double path for the flux outside the coils.

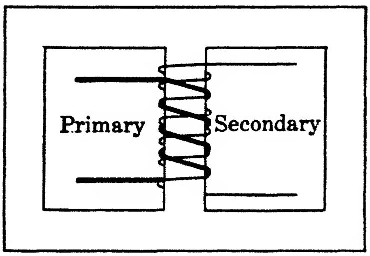


FIG. 249

Let us first investigate the properties of an *ideal transformer*, in which there are no energy losses and no leakage of flux, every tube of induction linking every turn of both primary and secondary, and in which the self-inductance of both windings is so large as to be effectively infinite.

As any flux in the core links all the turns of both coils, an induced e.m.f. in one is always accompanied by an induced e.m.f. in the other, and the amplitudes of these e.m.f.'s are in the same ratio as the number of turns in the coils. With finite e.m.f.'s the flux must evidently be vanishingly small since the inductances of the windings are very great. Due to this last condition a current in one winding must always be accompanied by a current in the other such that there is no appreciable amount of flux produced. This requires that the amplitudes of the currents are in the inverse ratio of the number of turns in the windings.

Thus, if $\mathcal{E}_p$ and $\mathcal{E}_s$ are the *terminal voltages* across the primary and secondary, respectively, and i_p and i_s are the currents, we have for the amplitudes

$$\left.\begin{aligned} \frac{\mathcal{E}_{p0}}{\mathcal{E}_{s0}} &= \frac{1}{a}, \\ \frac{i_{p0}}{i_{s0}} &= a, \end{aligned}\right\} \quad (117\text{–}1)$$

where a is the ratio of the number of turns in the secondary to the number in the primary.

Suppose we couple two circuits having the resistances and reactances indicated in Fig. 250 by means of an ideal transformer. We shall assume *positive coupling.* That is, having chosen

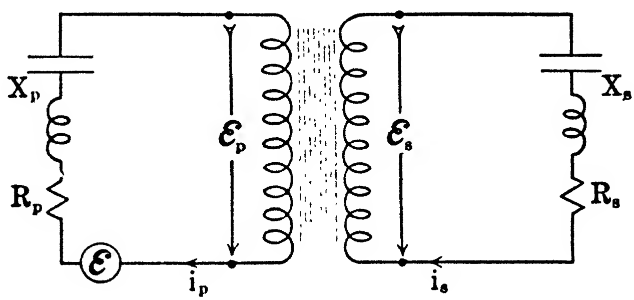


FIG. 250

positive directions for the currents in the circuits, we connect the transformer in such a way that the flux in the core links the two windings in the same sense (mutual inductance positive). Then the simultaneous induced e.m.f.'s in the windings are in the same sense. However, the terminal voltages are in opposite senses, for that in the primary is measured as a potential drop (p. 336) while that in the secondary represents an applied e.m.f. Thus the transformer equations are

$$\left.\begin{aligned} \mathcal{E}_p &= -\frac{1}{a}\mathcal{E}_s, \\ i_p &= -a i_s. \end{aligned}\right\} \quad (117\text{–}2)$$

In order to express the voltages and currents in terms of the impressed electromotive force $\mathcal{E}$, consider the impedance of the secondary circuit. The true impedance, exclusive of the transformer, is given by $Z_s = \mathcal{E}_{s0}/i_{s0}$. On the other hand, the equivalent impedance introduced into the primary through the transformer is $Z_e = \mathcal{E}_{p0}/i_{p0} = Z_s/a^2$, using (117–1). Therefore the characteristics of the primary circuit are the same as if the transformer was entirely removed and a resistance R_s/a^2 and a reactance X_s/a^2 were put in its place. This arrangement is equivalent to a simple series circuit with

$$\hat{R}_p = R_p + \frac{R_s}{a^2},$$

$$\hat{X}_p = X_p + \frac{X_s}{a^2},$$

for which calculations are easily made. Having determined $\mathcal{E}_p$ and i_p, $\mathcal{E}_s$ and i_s are obtained at once by means of (117–2). The property of a transformer of effectively changing an impedance is of great value as it allows us to adapt apparatus and circuits to use with other apparatus and circuits of entirely different impedances.

A well-constructed transformer in practice approximates an ideal transformer but it has some losses due to resistance of the windings and to hysteresis and eddy currents in the core. Also there is some leakage of flux in both windings and the inductance of the windings, although large, is finite. These defects may be represented in a circuit diagram by a fictitious ideal transformer of the given turn ratio with a small resistance and reactance in series with each winding and with a conductance and a susceptance shunted across the primary.

Two circuits coupled by a real transformer are shown in Fig. 251. The terminal voltages are $\mathcal{E}_p$ and $\mathcal{E}_s$, as before, and the total currents are i_p and i_s. Primed quantities refer to the fictitious ideal transformer. The current i_l flowing through the conductance g_n represents the core losses, while i_N flowing through

the susceptance b_n is the magnetizing current that produces the flux N_c in the core. The sum of i_l and i_N is the *exciting current* i_n, which added to i_p' gives the total primary current i_p. The

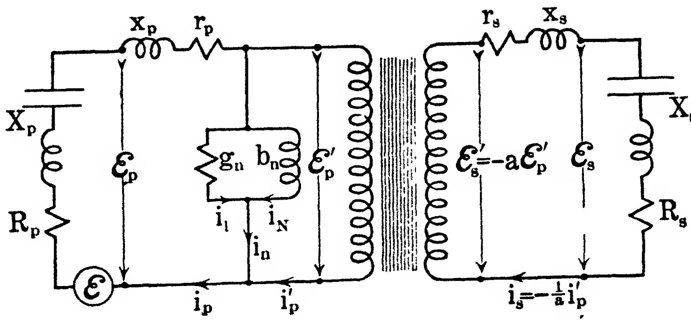


FIG. 251

quantities r_p and r_s are the primary and secondary winding resistances; x_p and x_s are the primary and secondary *leakage reactances*. The relations of the various currents and e.m.f.'s

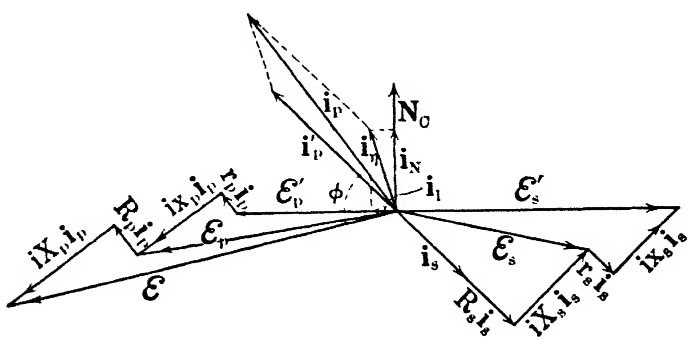


FIG. 252

are shown in the vector diagram, Fig. 252, which the reader should study carefully as it is an excellent example of graphical representation.

Problem 117a. An a.c. generator with a resistance of 10 ohms and negligible reactance is coupled to a load of 1000 ohms by means of an ideal transformer. In order to deliver maximum power to the load what turn ratio must the transformer have? Ans. 10.

Problem 117b. Referring to Fig. 250, the primary circuit contains a resistance of 60 ohms and an inductance of $1/\pi$ henry while the secondary contains a resistance of 900 ohms and an inductance of $3/\pi$ henry. The turn ratio of the transformer is 3. If the applied e.m.f. has an amplitude of 100 volts and a frequency of 60 cycles, find the primary and secondary currents (amplitude and phase). Ans. 0.442 amp, $-\pi/4$ radian; 0.147 amp, $3\pi/4$ radian.

Problem 117c. The circuits of the preceding problem are tuned to resonance, first by placing a condenser of appropriate size in the primary, then by placing one in the secondary. Find the ratio of the capacities. Ans. 9.

CHAPTER XIV

MEASUREMENTS WITH VARYING CURRENTS

118. Types and Methods of Measurement. — There are three fundamental quantities, resistance, inductance and capacity, to be determined in circuits carrying varying currents. Pure ohmic resistance is independent of frequency, except at very high frequencies, and it may be measured by d.c. methods when these are available. A special method for measuring very high resistance is given below. An inductance coil with an iron core has an apparent resistance, in addition to its ohmic resistance, due to hysteresis and eddy current losses. Similarly a condenser with a material dielectric has an apparent resistance due to dielectric losses (art. 20). These apparent resistances, which are often quite important, are usually measured in a.c. bridges simultaneously with the inductance or capacity with which they are associated.

Inductance and capacity are measured in a.c. bridges or by some special method, usually of a ballistic nature, depending on equipment available, accuracy required, and so forth. Measurements of inductance and capacity, except at high frequencies where resonance methods are used, fall into two general classes; (*a*) *comparison*, where the unknown quantity is determined in terms of known inductances and capacities, and (*b*) *absolute*, where the unknown is found in terms of some basic quantity, such as resistance. All a.c. bridge measurements described in this chapter are of the first class.

119. Measurement of High Resistance. — In article 88 we saw that if a condenser of capacity C with an initial charge q_0 is allowed to discharge through a resistance R the charge remaining after a time t is given by

$$q = q_0 e^{-(1/RC)t}. \tag{119-1}$$

This provides us with a method for measuring high resistance. With the aid of a ballistic galvanometer G the ratio q_0/q corresponding to any value of t may be found. This determines the value of RC experimentally, and, if C is known, we can calculate R. The method is usually limited to resistances of 10 megohms and above by the size of available condensers, for RC must have a value of 10 seconds or more to give accurate results.

The experimental arrangements are shown in Fig. 253. The known capacity is shunted by the resistance to be measured. A two-position key K connects the condenser to a steady e.m.f. $\mathcal{E}$ of a few volts or to the ballistic galvanometer, as required. It must be possible to leave the arm of this key midway between the contacts while the condenser is discharging through R.

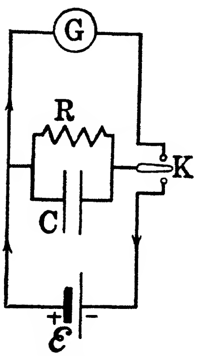


FIG. 253

A set of observations is begun by observing the galvanometer throw when K is snapped quickly from the lower to the upper contact. This throw corresponds to q_0 and should be made approximately full scale by proper choice of $\mathcal{E}$. The next observation is made by moving the arm of the key from the lower contact to the mid-position, allowing it to remain there for a measured time t, and then moving it to the upper position. The throw now corresponds to the charge q remaining after the condenser has discharged for a time t through R. Other observations are made with successively greater intervals of discharge, until the throw is too small to be accurately determined.

Now, taking the ratio of the first throw to each of the following, we have the values of q_0/q corresponding to various values of t. We may calculate RC for each case and take the mean, but it is simpler to find the average value of RC graphically. Taking the natural logarithm of (119–1),

$$t = RC \log \frac{q_0}{q}.$$

Thus, plotting t against $\log(q_0/q)$ gives a straight line whose slope is RC. Having determined RC we divide by the known value of C to obtain R.

The value of R just obtained applies actually to the parallel combination of the resistance to be measured and the insulation resistance of the condenser. The effect of the latter may not be entirely negligible, especially in measuring very high resistances, so we make another set of observations with the resistance element removed. The former value of R may then be corrected, if necessary, by means of the ordinary formula for resistances in parallel.

In the measurements just described it is essential to use a condenser entirely free from dielectric absorption, a condition usually satisfied by mica condensers. Also, the key employed must have extremely good insulation between its contacts to avoid leakage from the source of potential.

120. A. C. Bridges. — A number of a.c. bridge networks adapted to various measurements have been devised. Except under special conditions, they all involve a double balance. Usually one balance, called the d.c. balance, can be made with a steady e.m.f. and a D'Arsonval galvanometer. The remaining a.c. balance is then established with an alternating e.m.f. and a vibration galvanometer or telephone receiver, depending on the frequency. In practice double-pole double-throw switches in the battery and galvanometer arms are used to facilitate change from d.c. to a.c. condition and *vice versa*. However, these switching arrangements are not shown in the circuit diagrams of bridges described in this article for the sake of simplicity. The e.m.f. is designated in general by $\mathcal{E}$ and the appropriate current indicator by G. In this connection it is interesting to note that if the a.c. balance is independent of frequency, as is almost always the case, it holds for any wave form of applied e.m.f., since the latter may be expressed as a sum of sinusoidal components. Thus the a.c. balance may be obtained with the same steady e.m.f. and D'Arsonval galvanometer as the d.c. balance by

merely opening and closing a switch in the battery circuit. This procedure does not yield as accurate results as the other, however.

Certain precautions must be observed in the use of a.c. bridges. In order to obtain a sharp a.c. balance care must be taken that no random e.m.f.'s are introduced into the bridge. That is, the source of the alternating e.m.f. should be shielded, or located at some distance from the bridge, so that there is no magnetic coupling between the two. Inductive and capacitative coupling between different elements of the bridge should be reduced to a minimum by a careful arrangement of the parts. Capacity effects become increasingly important with higher frequencies. At frequencies where a telephone receiver is used it is advisable to couple the latter to the bridge by means of a transformer with some point on the primary grounded, in order to eliminate the effect of the observer's capacity. It is essential, of course, to employ only non-inductive resistance boxes designed for a.c. use.

Self-Inductance Bridge. — This bridge (Fig. 254) has been analyzed in article 114. The conditions for balance are

$$\left.\begin{aligned} \frac{R_1}{R_2} &= \frac{R_3}{R_4}, \\ \frac{L_1}{R_2} &= \frac{L_3}{R_4}. \end{aligned}\right\} \quad (120\text{–}1)$$

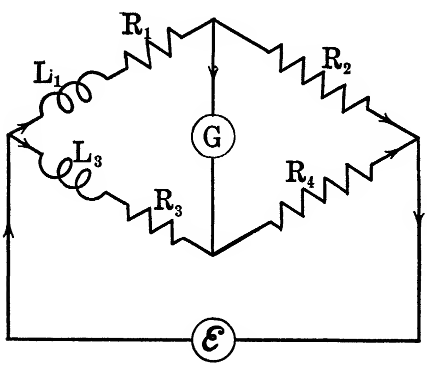


FIG. 254

In order to minimize the effect of any small residual inductance in the resistance boxes and to avoid insensitive arrangements of the bridge, it is desirable to have R_2 and R_4 at least of the same order of magnitude. To establish the d.c. balance under this condition it may be necessary to include a small external resistance in series with one of the inductances. For the a.c. balance L_1 and L_3 must now also be of the same order of magnitude.

Satisfactory use of this bridge is therefore limited to the measurement of inductances which do not differ greatly in magnitude from the extreme values at which the adjustable inductance L_3 can be set.

The accuracy of the results depends directly on the variable inductance or *inductometer* used. There are two standard types, the *Brooks-Weaver*, in which several coils connected in series move relative to one another in parallel planes, and the *Ayrton-Perry*, in which one coil rotates inside another.

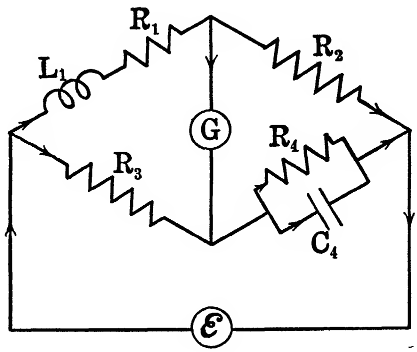


FIG. 255

Maxwell's L∼C Bridge. — A bridge for comparing a self-inductance and a capacity, due to Maxwell, is shown in Fig. 255. To find the conditions of balance we observe that, as in article 114,

$$\frac{Z_1}{Z_2} = \frac{Z_3}{Z_4},$$

that is,

$$\frac{R_1 + iL_1\omega}{R_2} = \frac{R_3}{\dfrac{R_4}{1 + iR_4C\omega}}.$$

Equating real parts and imaginary parts gives

$$\left.\begin{aligned} \frac{R_1}{R_2} &= \frac{R_3}{R_4}, \\ \frac{L_1}{R_2} &= R_3C_4. \end{aligned}\right\} \quad (120\text{–}2)$$

If either the self-inductance or the capacity is variable, the d.c. balance may be made in the usual way and left undisturbed while the a.c. balance is established. As variable capacities are

usually not available except in small sizes, Maxwell's bridge is adapted to the measurement of a capacity in terms of a variable inductance, or to the calibration of a variable inductance in terms of a capacity. If the inductance and the capacity are both fixed the bridge can be balanced only by the laborious process of trying different d.c. balances until one is found which is also an a.c. balance. Under these circumstances some other more suitable bridge should be used for the measurement.

Anderson's L∼C Bridge. — This is a modification of Maxwell's bridge which avoids the difficulties mentioned in the last paragraph. The capacity C_4 (Fig. 256) is now bridged across R_4 and an adjustable resistance R in the galvanometer arm, instead of across R_4 alone.

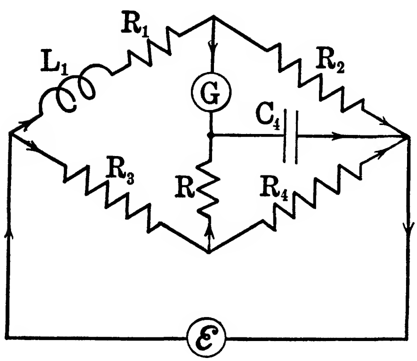


FIG. 256

Let us apply Kirchhoff's laws to this network. Denoting the complex currents in the main arms of the bridge by $\mathbf{i}_1$, $\mathbf{i}_2$, $\mathbf{i}_3$ and $\mathbf{i}_4$, and those flowing through R and C_4 by $\mathbf{i}_R$ and $\mathbf{i}_C$, respectively, we have at balance the special relations

$$(R_1 + \mathrm{i}L_1\omega)\mathbf{i}_1 = R_3\mathbf{i}_3 + R\mathbf{i}_R,$$

$$R_2\mathbf{i}_2 = -\,\mathrm{i}\frac{1}{C_4\omega}\mathbf{i}_C,$$

together with

$$\mathbf{i}_1 = \mathbf{i}_2, \qquad \mathbf{i}_R = \mathbf{i}_C.$$

Dividing the first equation by the second and using the current relations gives

$$\frac{R_1 + \mathrm{i}L_1\omega}{R_2} = \frac{R_3\left(\dfrac{\mathbf{i}_3}{\mathbf{i}_C}\right) + R}{-\,\mathrm{i}\dfrac{1}{C_4\omega}}. \qquad (120\text{–}3)$$

To find $\mathbf{i}_3/\mathbf{i}_C$ we note the general relations

$$R\mathbf{i}_R - i\frac{1}{C_4\omega}\mathbf{i}_C = R_4\mathbf{i}_4,$$
$$\mathbf{i}_3 = \mathbf{i}_R + \mathbf{i}_4.$$

With $\mathbf{i}_R = \mathbf{i}_C$ these reduce to

$$\left(R - i\frac{1}{C_4\omega}\right)\mathbf{i}_C = R_4\mathbf{i}_4,$$
$$\mathbf{i}_3 = \mathbf{i}_C + \mathbf{i}_4,$$

from which

$$\frac{\mathbf{i}_3}{\mathbf{i}_C} = \frac{R_4 + R - i\frac{1}{C_4\omega}}{R_4}.$$

Combining this with (120–3), we find

$$\frac{R_1 + iL_1\omega}{R_2} = \frac{\frac{R_3}{R_4}\left(R_4 + R - i\frac{1}{C_4\omega}\right) + R}{-i\frac{1}{C_4\omega}},$$

and the balance conditions are

$$\left.\begin{aligned} \frac{R_1}{R_2} &= \frac{R_3}{R_4}, \\ \frac{L_1}{R_2} &= C_4\left[\frac{R_3}{R_4}(R_4 + R) + R\right]. \end{aligned}\right\} \quad (120\text{–}4)$$

After the usual d.c. adjustment has been made the a.c. balance is established by means of R, so neither inductance nor capacity need be variable. It is usually desirable to have R_3 and R_4 of the same order of magnitude. Often they are made equal, in which case

$$L_1 = C_4R_2(2R + R_4).$$

The Anderson bridge does not suffer any serious restrictions, being capable of accurate measurements over a wide range of values. Its characteristics are described in detail by Rosa and Grover * who have used it in standardization work.

* Bulletin of the Bureau of Standards, Vol. I, p. 291.

Owen's L∼C Bridge. — This differs from the two preceding bridges in that two capacities (Fig. 257) are involved, an adjustable resistance R_3 being placed in series with one of them.

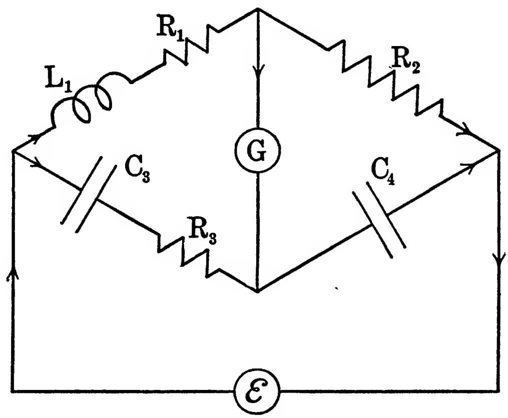


FIG. 257

The balance conditions are obtained from

$$\frac{Z_1}{Z_2} = \frac{Z_3}{Z_4},$$

that is, from

$$\frac{R_1 + iL_1\omega}{R_2} = \frac{R_3 - i\dfrac{1}{C_3\omega}}{-i\dfrac{1}{C_4\omega}}.$$

Evidently we must have

$$\left.\begin{aligned} \frac{R_1}{R_2} &= \frac{C_4}{C_3}, \\ \frac{L_1}{R_2} &= C_4R_3. \end{aligned}\right\} \quad (120\text{–}5)$$

The first condition does not represent an ordinary Wheatstone bridge balance, since no d.c. current flows through the condensers. However, the balance can be made by d.c. means. Suppose that the galvanometer arm is opened and a steady e.m.f. is applied to the bridge. If $R_1/R_2 = C_4/C_3$ the ends of the

galvanometer arm are at the same potential, and the galvanometer switch may be closed without effect. On the other hand, if the balance equation is not satisfied, the ends of the galvanometer arm assume different potentials. Closing the galvanometer switch now results in a ballistic throw. The d.c. balance is thus established by varying one of the bridge elements, usually R_2, until there is no throw of the galvanometer when its switch is closed. Care must be taken to discharge the condensers after each observation. The a.c. balance is obtained by adjusting R_3, the other balance thereby remaining undisturbed.

Since L_1 is proportional to R_3 the bridge is useful for measuring inductances of any size. It is not as accurate as the Anderson bridge but is easy to assemble and simple to operate. It is necessary to use good condensers whose apparent resistance is negligible, if a sharp balance is to be obtained. Usually it is advantageous to have C_3 and C_4 of the same order of magnitude, which may require some external resistance in series with L_1.

Mutual Inductance Bridge. — A simple network for the comparison of mutual inductances is shown in Fig. 258. The equa-

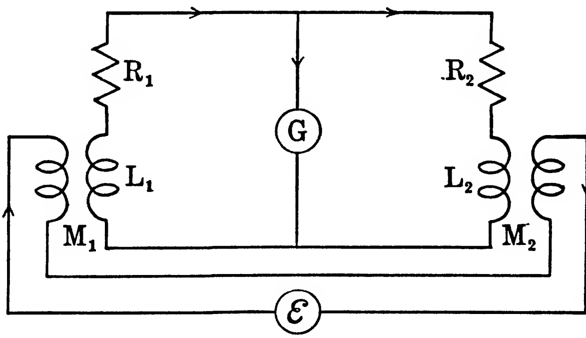


FIG. 258

tions for balance are readily found with the aid of Kirchhoff's laws. Denoting the currents in the bridge arms by $\mathbf{i}_1$ and $\mathbf{i}_2$, and the current in the generator circuit by $\mathbf{i}$, we must have

$$(R_1 + \mathrm{i}L_1\omega)\mathbf{i}_1 + \mathrm{i}M_1\omega\mathbf{i} = 0,$$
$$(R_2 + \mathrm{i}L_2\omega)\mathbf{i}_2 + \mathrm{i}M_2\omega\mathbf{i} = 0,$$

with $\mathbf{i}_1 = \mathbf{i}_2$. By division we obtain

$$\frac{R_1 + \mathrm{i}L_1\omega}{R_2 + \mathrm{i}L_2\omega} = \frac{\mathrm{i}M_1\omega}{\mathrm{i}M_2\omega},$$

which gives

$$\left.\begin{aligned} \frac{M_1}{R_1} &= \frac{M_2}{R_2}, \\ \frac{M_1}{L_1} &= \frac{M_2}{L_2}. \end{aligned}\right\} \quad (120\text{–}6)$$

Both equations represent a.c. conditions, so the two balances must be made simultaneously. Fortunately this is not as laborious as it might appear, for by making the resistances large the effect of the self-inductances is made small and the bridge can be approximately balanced without regard to the ratio L_1/L_2. The resistances are now reduced without changing their ratio, and the approximate balance improved by an adjustment of L_1/L_2. Hereafter slight readjustments of R_1/R_2 and L_1/L_2, made alternately, soon establish a complete balance. Note that the mutual inductances must be connected in the same sense to make this balance possible.

As ordinarily used the bridge consists of two fixed mutual inductances, one of which is known, with adjustable resistances in series. A variable, but not necessarily calibrated, self-inductance is placed in whichever arm of the bridge the self-inductance associated with the mutual inductance is too small to satisfy the balance condition. Since R_1 and R_2 are the total resistances of the bridge arms, the resistances of the coils must be found and added to the adjustable resistances.

Maxwell's $M{\sim}L$ Bridge. — The comparison of a mutual inductance and the self-inductance of one of its coils is most simply made by means of the bridge illustrated in Fig. 259.

Analysis of the network is carried out in the usual manner. Distinguishing currents in the bridge arms by numerical subscripts and denoting the currents flowing through M and R by

$\mathbf{i}_M$ and $\mathbf{i}_R$, respectively, we have under the conditions of balance

$$(R_1 + iL_1\omega)\mathbf{i}_1 + iM\omega\mathbf{i}_M = R_3\mathbf{i}_3,$$
$$R_2\mathbf{i}_2 = R_4\mathbf{i}_4,$$

together with

$$\mathbf{i}_1 = \mathbf{i}_2, \qquad \mathbf{i}_3 = \mathbf{i}_4.$$

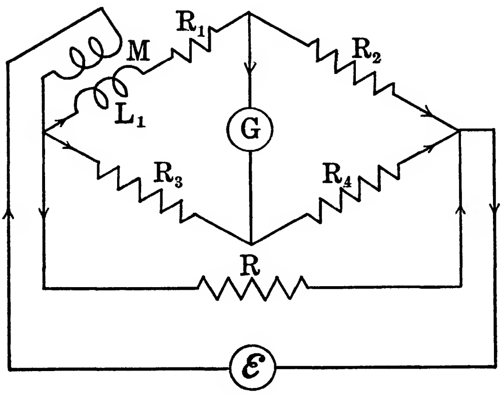


FIG. 259

In addition there are the general relations

$$R_3\mathbf{i}_3 + R_4\mathbf{i}_4 = R\mathbf{i}_R,$$
$$\mathbf{i}_1 + \mathbf{i}_3 + \mathbf{i}_R = \mathbf{i}_M.$$

Using $\mathbf{i}_3 = \mathbf{i}_4$, and eliminating $\mathbf{i}_R$, these give

$$\mathbf{i}_1 + \left(1 + \frac{R_3 + R_4}{R}\right)\mathbf{i}_3 = \mathbf{i}_M.$$

We can now eliminate $\mathbf{i}_M$ from the first condition equation and divide by the second, obtaining

$$\frac{(R_1 + iL_1\omega) + iM\omega}{R_2} = \frac{R_3 - iM\omega\left(1 + \dfrac{R_3 + R_4}{R}\right)}{R_4}.$$

The balance equations are thus

$$\left.\begin{aligned}\frac{R_1}{R_2} &= \frac{R_3}{R_4},\\ \frac{L_1 + M}{R_2} &= \frac{-M\left(1 + \dfrac{R_3 + R_4}{R}\right)}{R_4}.\end{aligned}\right\} \quad (120\text{–}7)$$

The a.c. balance is made by adjustment of R, which does not disturb the previously obtained d.c. balance. Rearrangement of the a.c. balance equation gives

$$M = -L_1 \frac{R_4}{R_2\left(1 + \dfrac{R_3}{R}\right) + R_4\left(1 + \dfrac{R_2}{R}\right)},$$

which shows that the magnitude of L_1 must be greater than that of M. This is a distinct limitation on the usefulness of the bridge. The negative sign signifies, of course, that M must be connected to give negative coupling, so that the induced e.m.f. due to L_1 is opposed by that due to M.

Heaviside's M~L Bridge. — This bridge (Fig. 260) is more

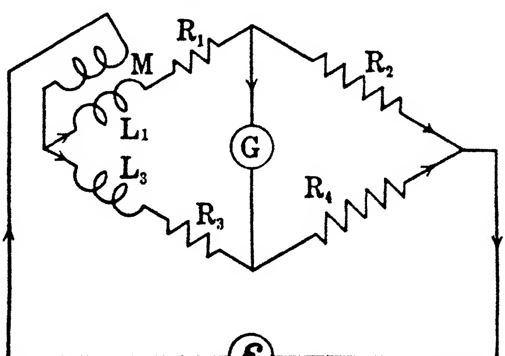


FIG. 260

sensitive than Maxwell's and it does not suffer from restrictions on the relative magnitudes of M and L_1. However, it involves a self-inductance L_3 in addition to that associated with the

mutual inductance, and either the independent self-inductance or the mutual inductance must be adjustable.

Using the same notation as in the case of Maxwell's bridge just preceding, the condition equations at balance are

$$(R_1 + \mathrm{i}L_1\omega)\mathbf{i}_1 + \mathrm{i}M\omega\mathbf{i}_M = (R_3 + \mathrm{i}L_3\omega)\mathbf{i}_3,$$

$$R_2\mathbf{i}_2 = R_4\mathbf{i}_4,$$

with

$$\mathbf{i}_1 = \mathbf{i}_2, \qquad \mathbf{i}_3 = \mathbf{i}_4.$$

Eliminating $\mathbf{i}_M$ from the first condition equation by means of the general relation $\mathbf{i}_1 + \mathbf{i}_3 = \mathbf{i}_M$, and dividing by the second condition equation, we have

$$\frac{(R_1 + \mathrm{i}L_1\omega) + \mathrm{i}M\omega}{R_2} = \frac{(R_3 + \mathrm{i}L_3\omega) - \mathrm{i}M\omega}{R_4}.$$

The balance conditions are therefore

$$\left.\begin{aligned} \frac{R_1}{R_2} &= \frac{R_3}{R_4}, \\ \frac{L_1 + M}{R_2} &= \frac{L_3 - M}{R_4}. \end{aligned}\right\} \qquad (120\text{–}8)$$

The d.c. balance is as usual. The a.c. balance is established by adjusting L_3 or M. In order to find either of these quantities in terms of the other L_1 must be known, of course. Solving the a.c. balance equation for M gives

$$M = \frac{L_3R_2 - L_1R_4}{R_2 + R_4},$$

which shows that M must be connected for positive coupling or for negative depending on whether L_3R_2 is greater than L_1R_4 or less. If we use an adjustable resistance in series with L_1 or L_3, as required, to establish the d.c. balance, we may set $R_2 = R_4$. This gives

$$M = \tfrac{1}{2}(L_3 - L_1),$$

a very convenient relation.

Incidentally, the Heaviside bridge may be used differentially to measure self-inductance in terms of mutual inductance alone. Thus, suppose M is a variable calibrated mutual inductance, while L_1 and L_3 are fixed self-inductances, preferably of the same order of magnitude. First, let the bridge be balanced with $R_2 = R_4$, so that $M' = \frac{1}{2}(L_3 - L_1)$. Then, having placed the inductance L to be measured in series with L_3, let the bridge be rebalanced. This time we must have

$$M'' = \tfrac{1}{2}($$

Taking the difference of these equations, we see that

$$L = 2(M'' - M'),$$

where M' and M'' represent the two settings of the variable mutual inductance. In addition to simplicity, this method of measuring self-inductance has the great advantage that errors due to residual inductance in R_2 and R_4, capacity effects and so forth, affect M' and M'' equally and so eliminate themselves.

Heydweiller's $M \sim C$ Bridge. — A convenient method for comparing a mutual inductance and a capacity is illustrated in

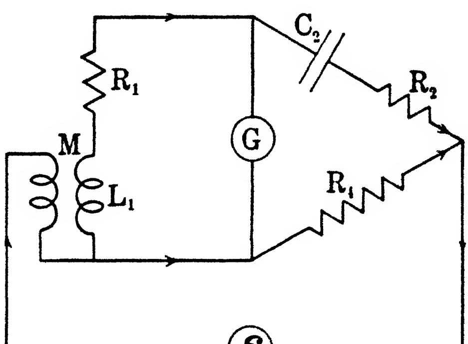


Fig. 261

Fig. 261. Both the inductance and the capacity must have adjustable resistances in series with them, but no other variable elements are required.

When the bridge is balanced we have

with

where the notation is the same as that used for the two preceding bridges. Eliminating $\mathbf{i}_M$ from the first equation by means of the general relation $\mathbf{i}_1 + \mathbf{i}_4 = \mathbf{i}_M$, and dividing by the second equation, we obtain

The corresponding balance equations are

$$\left.\begin{aligned} & R_2 + R_4, \\ M &= -\,C_4R_1R_4, \end{aligned}\right\} \quad (120\text{–}9)$$

the negative signs indicating the connection of the mutual inductance for negative coupling, as in the case of the Maxwell bridge. The balances are both a.c. and must be made simultaneously. This is done simply by adjusting R_1 and R_2 alternately, since each of these appears in only one of the balance equations. Note that L_1 must be greater in magnitude than M. If necessary a separate self-inductance can be placed in series with the appropriate coil of the mutual inductance to establish this condition.

Capacity Bridge. — This bridge (Fig. 262) compares two capacities and, incidentally, establishes a relation between their apparent resistances. When the standard capacity is supplied by a good condenser with negligible resistance, the unknown capacity and its associated resistance are measured simultaneously, which is very convenient.

As in other bridges of the Wheatstone type, the balance con-

ditions are obtained from

$\overline{Z}_4$.

In this case we have

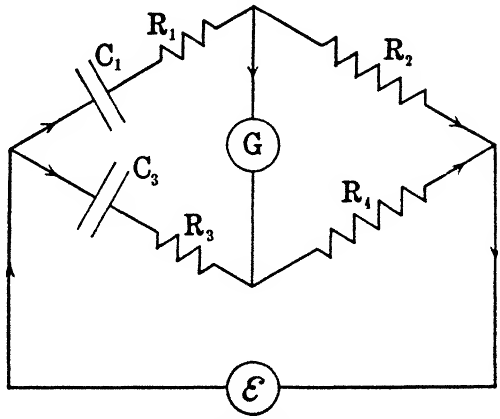


FIG. 262

$$\frac{1}{C} \qquad R_3 - i\,\frac{1}{C}$$

and the balance equations are

(120–10)

As the bridge is ordinarily used, C_1 is the unknown capacity and R_1 is its associated apparent resistance, while C_3 is the standard capacity, with an adjustable resistance R_3 in series. No d.c. current can flow, so the balances must be made simultaneously with a.c. C_3 and R_3 are varied alternately, beginning with C_3, until a complete balance is found. If C_3 is fixed, the capacity balance is made by means of R_2 or R_4. When possible R_2 and R_4 should be of the same order of magnitude to minimize the effect of residual inductance.

From R_1 and C_1 we may determine the apparent power factor (art. 90) of the condenser. Thus,

$$\cos \phi_1 = \frac{1}{\sqrt{1 + \tan^2 \phi_1}} = \frac{1}{\sqrt{1 + \dfrac{1}{R_1^2 C_1^2 \omega^2}}},$$

where $\omega/2\pi$ is the frequency at which the bridge is balanced. Actually, with any usable condenser, $R_1 C_1 \omega$ is very small compared to unity, so that, effectively,

$$\cos \phi_1 = R_1 C_1 \omega. \tag{120-11}$$

The power factor of a condenser is important because it is an excellent criterion of quality, the smaller the power factor, the better the condenser, in general. Experiment shows that the power factor is practically independent of frequency, at least over moderate ranges of frequency variation.

Frequency Bridge. — Any bridge in which the balance depends on frequency may be used to measure frequency. A simple but effective bridge for this purpose is shown in Fig. 263.

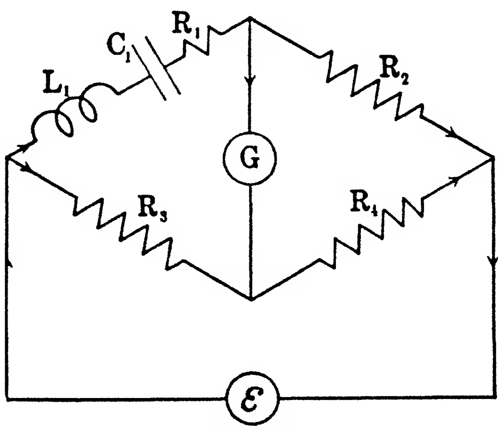


FIG. 263

Evidently the balance conditions are

$$R_1 \quad R_3$$

$$\Bigg\} \quad (120-12)$$

$$= 0.$$

Using a variable inductance or capacity, the scale may be calibrated to read frequency directly, since the a.c. balance is independent of the resistance ratios.

121. Absolute Measurements. — We have already seen (art. 105) that resistance may be determined in absolute measure. Determinations of inductance and capacity in which resistance is the only electrical quantity involved are therefore regarded as absolute measurements. Examples of this type of measurement are given below.

Self-Inductance. — The experimental arrangements for the absolute measurement of self-inductance (Fig. 264) consist

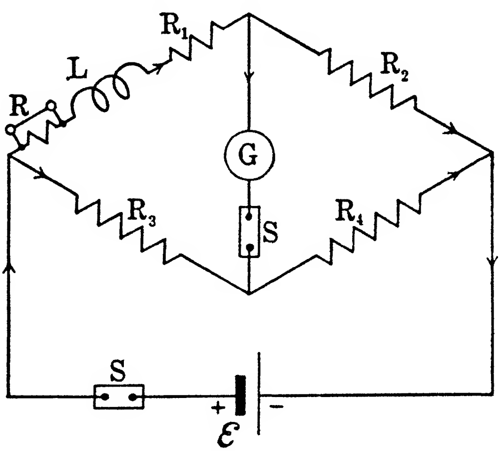


FIG. 264

essentially of a d.c. Wheatstone bridge with the unknown inductance L included in one of the arms. In the same arm is a very small resistance R which can be short-circuited by a copper rod dipping into mercury cups or by some other device with negligible contact resistance. Switches (S) are located as shown in the figure.

The galvanometer must have sufficiently small damping (including the effect of the external circuit) to allow ballistic

observations to be made as well as steady current measurements. A moving needle galvanometer is preferable but a D'Arsonval galvanometer may be used if the galvanometer circuit is opened immediately after the ballistic impulse has been given. The resistance elements must have negligible self-inductance as compared with L.

To measure L, an ordinary d.c. balance is established and then, with the galvanometer circuit still closed, the battery circuit is opened. The current i through L falls from its steady value i_1 to zero in a very short interval of time during which there is an induced e.m.f. $-L\dfrac{di}{dt}$. This e.m.f. produces a momentary current through the galvanometer, whose ballistic throw is proportional to the total charge passing. Neglecting the effect of the inductance of the galvanometer for the moment, the current through G must be proportional to the e.m.f. at each instant and the total charge is given by

$$\int_{i_1}^{0} k\left(-L\frac{di}{dt}\right)dt = -kL\int_{i_1}^{0} di = kLi_1, \qquad (121\text{–}1)$$

where k is a constant of proportionality. Now, as is illustrated for a special case in article 98, the charge passing through the galvanometer due to changing flux is independent of its self-inductance, so (121–1) gives the charge correctly regardless of the magnitude of the inductance neglected. Expressing the charge in terms of the first throw α_1, we have

$$kLi_1 = \frac{P_0K_i}{2\pi}\left(1+\frac{\lambda}{2}\right)\alpha_1 \qquad (121\text{–}2)$$

from (98–17), where P_0 is the period, λ is the half logarithmic decrement and K_i is the galvanometer constant.

Next, with battery and galvanometer circuits both closed and the previous d.c. balance unchanged, the resistance R is short-circuited. This has the same effect as introducing an e.m.f. equal to Ri_1 and causes a steady current kRi_1 to flow through the galvanometer. We use the undisturbed value i_1 for the current

because R_1 is always taken so large compared to R that the current in this arm of the bridge is practically unaffected. The galvanometer now has a steady deflection α given by

$$kRi_1 = K_i\alpha \tag{121-3}$$

according to (74-8).

Dividing (121-2) by (121-3) we have the desired absolute relation between L and R, namely,

$$\frac{L}{R} = \frac{P_0}{2\pi}\left(1 + \frac{\lambda}{2}\right)\frac{\alpha_1}{\alpha}. \tag{121-4}$$

P_0 and λ must be determined (art. 98) with the galvanometer under working conditions.

For accurate results R should be so chosen that α is of the same order of magnitude as α_1. It must be remembered that α_1/α is the ratio of the angular deflections. This is equal to the ratio of the corresponding scale readings if the scale is curved. If the scale is straight the ratio of the scale readings may require slight correction.

Mutual Inductance. — The absolute measurement of mutual inductance depends on the same principle as the corresponding measurement of self-inductance just described. Circuit arrangements are illustrated in Fig. 265. A double-pole double-throw switch S allows the galvanometer to be coupled to the battery circuit through the mutual inductance M alone or to be shunted across a very small resistance R. R_1 is the total resistance of the battery circuit exclusive of R. If variable, it may conveniently be used to control sensitivity, but it must always be large compared to R. R_2 is the total resistance of the galvanometer circuit. There is a key K as shown. As in the previous case the galvanometer should be of the

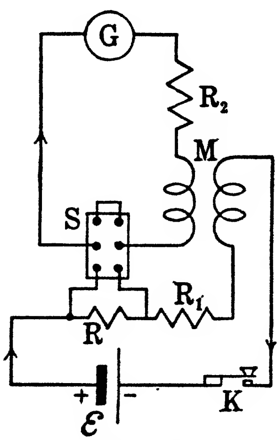


FIG. 265

moving needle type to avoid excessive damping, unless the circuit is opened immediately after a ballistic impulse.

To measure M, the switch S is placed in the "up" position with K closed. The battery circuit is then broken and the current i in it falls from the steady value i_1 to zero, inducing in the galvanometer circuit an e.m.f. $-M\frac{di}{dt}$. A momentary current flows through the galvanometer, whose resulting throw is proportional to the total charge passing. This charge, being independent of the self-inductance of the galvanometer (see p. 495), is given by

$$\int_{i_1}^{0}\frac{1}{R_2}\left(-M\frac{di}{dt}\right)dt = -\frac{M}{R_2}\int_{i_1}^{0}di = \frac{Mi_1}{R_2}.$$

Expressing the charge in terms of the first throw α_1 by means of (98–17), we have

$$\frac{Mi_1}{R_2} = \frac{P_0K_i}{2\pi}\left(1+\frac{\lambda}{2}\right)\alpha_1, \qquad (121\text{–}5)$$

where P_0 is the period, λ is the half logarithmic decrement and K_i is the galvanometer constant.

Next S is placed in the "down" position. With K closed there is a steady deflection α of the galvanometer corresponding to an e.m.f. Ri_1 in the galvanometer circuit. We use the same value i_1 as before for the current in the battery circuit, since R is so small compared to R_1 that the shunting effect of R_2 is negligible. The current through the galvanometer is evidently Ri_1/R_2 and hence, by (74–8),

$$\frac{Ri_1}{R_2} = K_i\alpha. \qquad (121\text{–}6)$$

Dividing (121–5) by (121–6) gives the absolute relation between M and R,

$$\frac{M}{R} = \frac{P_0}{2\pi}\left(1+\frac{\lambda}{2}\right)\frac{\alpha_1}{\alpha}. \qquad (121\text{–}7)$$

As in the case of the corresponding measurement of self-induc-

tance above, P_0 and λ must be determined with the galvanometer under working conditions, and R should be chosen so that α and α_1 are of the same order of magnitude.

Capacity. — This absolute measurement is very similar to the preceding one, to which the reader should refer. The unknown capacity C (Fig. 266) is charged to a potential difference $\mathcal{E}$ and then discharged through the galvanometer by means of a two-position key K. The first throw α_1 is therefore given by

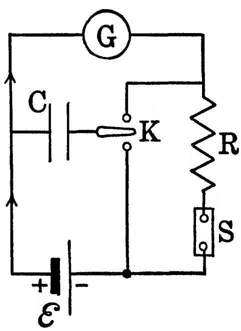


FIG. 266

$$C\mathcal{E} = \frac{P_0 K_i}{2\pi}\left(1 + \frac{\lambda}{2}\right)\alpha_1.$$

Next, the switch S is closed, allowing a steady current to flow through the galvanometer which now has a very high resistance R in series with it. The galvanometer deflection α is determined by

$$\frac{\mathcal{E}}{R} = K_i\alpha.$$

Dividing the charge equation by the current equation, we obtain

$$RC = \frac{P_0}{2\pi}\left(1 + \frac{\lambda}{2}\right)\frac{\alpha_1}{\alpha}. \qquad (121\text{–}8)$$

As in the other absolute measurements R should be chosen so that α and α_1 are of the same order of magnitude. Note that in this case the galvanometer operates ballistically on open circuit, so the D'Arsonval type may be employed to advantage.

The above method of determining C is not as accurate as the tuning fork method described in article 107, especially when C is small. It is simpler to employ, however, and does not require special apparatus.

122. Measurement of Large Self-Inductance. — Iron cored coils of large self-inductance, called choke coils (art. 90), are

often used to separate direct currents from alternating currents. Such coils permit free passage of direct current but offer a high impedance to alternating current. It is difficult to determine the inductance L of a choke coil by any of the ordinary methods described in this chapter, both because of its magnitude, which may run to a number of henries, and because of its dependence on the magnetic state of the iron core, that is, on the amount of direct current flowing simultaneously with the alternating current. To be of any value, measurements must be made on choke coils under the conditions in which the coils are to be used.

A simple method of measuring L under practical conditions has been devised by Turner.* The principle of the method is indicated in Fig. 267. The choke coil with a switch S in series

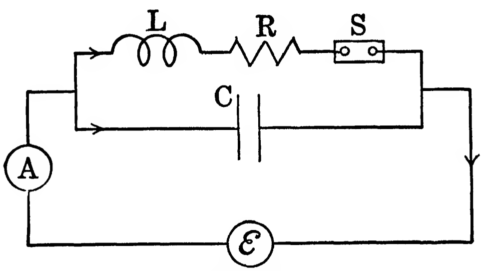


FIG. 267

is shunted by a variable capacity C to form a divided circuit. An a.c. ammeter A is included in the line to show the magnitude of the total current, which is proportional to the real admittance Y (art. 112) of the circuit.

Since the complex admittance is

$$\mathbf{Y} = \frac{1}{R + iL\omega} + iC\omega = \frac{R}{R^2 + L^2\omega^2} - i\frac{L\omega - C\omega(R^2 + L^2\omega^2)}{R^2 + L^2\omega^2},$$

we have

$$Y = \sqrt{\left[\frac{R}{R^2 + L^2\omega^2}\right]^2 + \left[\frac{L\omega - C\omega(R^2 + L^2\omega^2)}{R^2 + L^2\omega^2}\right]^2}$$

$$= \sqrt{\frac{1 - 2LC\omega^2}{R^2 + L^2\omega^2} + C^2\omega^2}. \tag{122-1}$$

* I. R. E., Vol. 16, p. 1559 (Nov. 1928).

From this it appears that if

$$1 - 2LC\omega^2 = 0, \tag{122-2}$$

the admittance Y and hence the magnitude of the line current have the same values as if the capacity alone was in the circuit. Therefore to determine L we have only to adjust C until opening and closing S does not affect the ammeter reading. Then (122–2) is satisfied and we can calculate L if C and ω are known.

In order to send direct current through the coil simultaneously with the alternating current, a battery may, of course, be introduced in series with the alternating e.m.f. This is, however, objectionable in practice because it prevents a sensitive ammeter from being used in the line. A better plan is to use two identical choke coils, if available, in parallel, with a battery and potentiometer arranged to produce a circulating direct current which does not pass into the line. Such an arrangement is shown

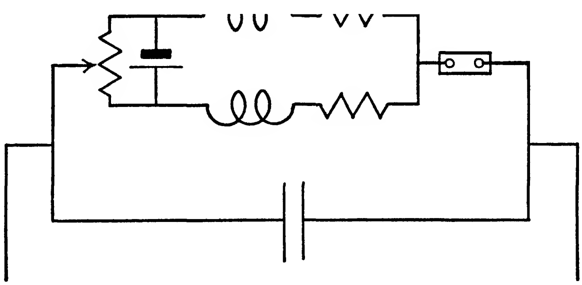


FIG. 268

schematically in Fig. 268. The inductance now measured is one-half that possessed by either coil alone.

In making the measurements described above, care must be taken that the change of phase which takes place when S is opened or closed does not produce a change of voltage across the divided circuit and hence a change of total current even when (122–2) is satisfied. This effect may be avoided by keeping the impedance of the ammeter and of the source of e.m.f. small.

123. Ballistic Comparison Measurements. — A ballistic galvanometer may be used to compare any two sources of charge, such as two mutual inductances, a mutual inductance and a capacity, or two capacities. The latter comparison is particularly useful, as it provides a very simple means of finding an unknown capacity in terms of a standard when measurement of the apparent resistance is not required.

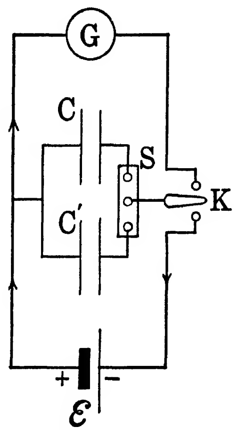


FIG. 269

The experimental arrangements (Fig. 269) consist of a single-pole double-throw switch S and a two-position key K so connected that either the unknown or the standard capacity may be charged to a potential difference $\mathcal{E}$ and then discharged through the galvanometer.

If C and C' are the capacities, and α_1 and α_1' are the corresponding first throws, we have, from article 98,

$$C\mathcal{E} = K_q'\alpha_1,$$
$$C'\mathcal{E} = K_q'\alpha_1'.$$

Hence, by division,

$$\alpha_1$$
$$C'$$

In making the observations the key should be thrown from charge to discharge quickly, in order to avoid loss of charge by leakage or by dielectric absorption.

CHAPTER XV

COUPLED CIRCUITS, FILTERS AND LINES

124. Types of Coupled Circuits. — It is often convenient to couple two or more simple circuits by means of inductance or capacity. The combination of circuits forms a network and may be treated as such. It is easier to obtain a physical insight into the behavior of the circuits, however, if their separate identities are maintained.

Confining ourselves to two circuits for the present, the simplest types of coupling are illustrated in Fig. 270. In (*a*) the coupling is *inductive*, in (*b*) *direct*, and in (*c*) *capacitative*.

The larger the fraction of the inductance or of the capacity of each circuit that is mutual, the more the circuits react on one another. The square root of the product of these fractions for any pair of coupled circuits of the type illustrated is called the *coefficient of coupling* of the circuits. For example, in case (*a*) the coefficient of coupling is given by

$$k = \sqrt{\left(\frac{\pm M}{L_1}\right)\left(\frac{\pm M}{L_2}\right)} = \frac{\pm M}{\sqrt{L_1 L_2}}, \qquad (124\text{–}1)$$

the positive sign being used when M is positive and the negative sign when M is negative, so that k is always positive. As the greatest possible magnitude of the mutual inductance is $\sqrt{L_1 L_2}$ we see that $0 \leqq k \leqq 1$. Similarly, in case (*b*),

$$k = \frac{L_M}{\sqrt{(L' + L_M)(L'' + L_M)}}, \qquad (124\text{–}2)$$

and in case (*c*),

$$k = \frac{\dfrac{1}{C_M}}{\sqrt{\left(\dfrac{1}{C'} + \dfrac{1}{C_M}\right)\left(\dfrac{1}{C''} + \dfrac{1}{C_M}\right)}}. \qquad (124\text{–}3)$$

The general behavior of all three types of coupled circuits is the same. Except under limiting conditions there are two characteristic frequencies, both for free oscillations and for forced oscillations. An investigation of inductively coupled circuits, the most common type, is carried out in the next two articles.

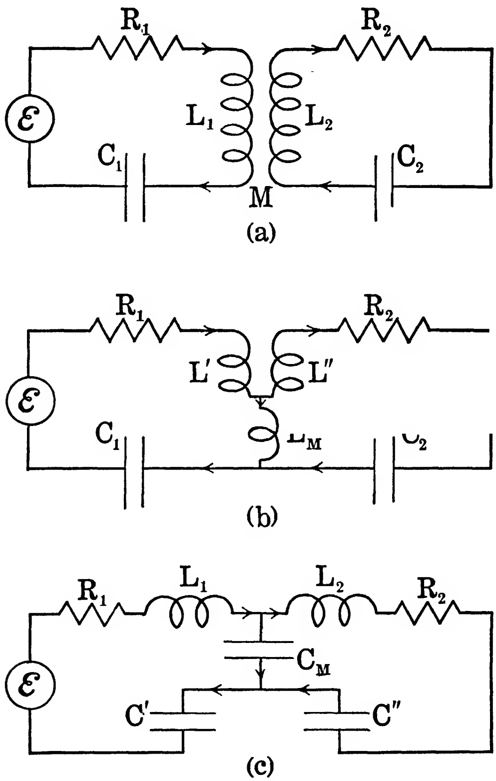


FIG. 270

Problem 124a. Show that direct coupling between two circuits may be replaced by inductive coupling such that

$$M = - \qquad L_1 = L' + \qquad = L''$$

without changing the electrical characteristics of the circuits.

Problem 124b. If C_1 and C_2 are the total series capacities in two capacitatively coupled circuits, show that the coefficient of coupling may be put in the form $\sqrt{C_1}$

125. Inductively Coupled Circuits Freely Oscillating. — Suppose an electrical impulse is given to the coupled circuits of Fig. 271 by some means, such as allowing the condenser C_1 in the primary circuit to discharge, there being no applied e.m.f. in either circuit. Then, if the resistances are not too great, free oscillations take place as in the case of a single circuit (art. 89).

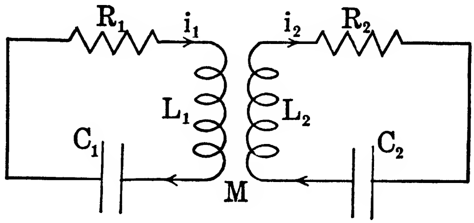


FIG. 271

The circuit equations are

$$\left.\begin{aligned} L_1\frac{di_1}{dt} + R_1i_1 + \frac{q_1}{C_1} &= -M\frac{di_2}{dt}, \\ L_2\frac{di_2}{dt} + R_2i_2 + \frac{q_2}{C_2} &= -M\frac{di_1}{dt}, \end{aligned}\right\} \qquad (125\text{-}1)$$

or, differentiating with respect to the time,

$$\left.\begin{aligned} L_1\frac{d^2i_1}{dt^2} + R_1\frac{di_1}{dt} + \frac{1}{C_1}i_1 &= -M\frac{d^2i_2}{dt^2}, \\ L_2\frac{d^2i_2}{dt^2} + R_2\frac{di_2}{dt} + \frac{1}{C_2}i_2 &= -M\frac{d^2i_1}{dt^2}. \end{aligned}\right\} \qquad (125\text{-}2)$$

The first step in solving these simultaneous differential equations is to express each in terms of one current only. Having multiplied either equation through by M and differentiated twice, we may substitute in the relation so obtained the value of the right-hand member of the other equation, getting

$$\begin{aligned} &(L_1L_2 - M^2)\frac{d^4i_1}{dt^4} + (R_1L_2 + R_2L_1)\frac{d^3i_1}{dt^3} \\ &+\left(\frac{L_1}{C_2} + R_1R_2 + \frac{L_2}{C_1}\right)\frac{d^2i_1}{dt^2} + \left(\frac{R_1}{C_2} + \frac{R_2}{C_1}\right)\frac{di_1}{dt} + \frac{1}{C_1C_2}i_1 = 0, \\ &(L_1L_2 - M^2)\frac{d^4i_2}{dt^4} + (R_1L_2 + R_2L_1)\frac{d^3i_2}{dt^3} \\ &+\left(\frac{L_1}{C_2} + R_1R_2 + \frac{L_2}{C_1}\right)\frac{d^2i_2}{dt^2} + \left(\frac{R_1}{C_2} + \frac{R_2}{C_1}\right)\frac{di_2}{dt} + \frac{1}{C_1C_2}i_2 = 0. \end{aligned} \qquad (125\text{-}3)$$

Since these equations are identical in form the currents must be of the same form. That is, damping constants and frequencies of oscillation (art. 89) must be the same in the two circuits. Exact solutions of (125–3) are extremely complicated. Fortunately in practical applications of oscillating circuits we always use circuits with small damping, so we shall restrict our analysis to this condition and obtain approximate solutions. Thus, referring to frequencies as zero order quantities, damping constants as first order, damping constants squared as second order and so on, we shall neglect a quantity of any order in comparison with one of lower order.

Let us divide both equations of (125–3) by L_1L_2. If we put

$$\alpha_1 \equiv \frac{R_1}{2L_1}, \qquad \alpha_2 \equiv \frac{R_2}{2L_2},$$

and

$$\omega_1 \equiv \frac{1}{\sqrt{L_1C_1}}, \qquad \omega_2 \equiv \frac{1}{\sqrt{L_2C_2}},$$

we obtain

$$\left.\begin{aligned}(1-k^2)\frac{d^4i_1}{dt^4} + 2(\alpha_1+\alpha_2)\frac{d^3i_1}{dt^3} + (\omega_1^2 + 4\alpha_1\alpha_2 + \omega_2^2)\frac{d^2i_1}{dt^2}&\\ + 2(\alpha_1\omega_2^2 + \alpha_2\omega_1^2)\frac{di_1}{dt} + \omega_1^2\omega_2^2 i_1 = 0,&\\ (1-k^2)\frac{d^4i_2}{dt^4} + 2(\alpha_1+\alpha_2)\frac{d^3i_2}{dt^2} + (\omega_1^2 + 4\alpha_1\alpha_2 + \omega_2^2)\frac{d^2i_2}{dt^2}&\\ + 2(\alpha_1\omega_2^2 + \alpha_2\omega_1^2)\frac{di_2}{dt} + \omega_1^2\omega_2^2 i_2 = 0,&\end{aligned}\right\} \quad (125\text{–}4)$$

where k is the coefficient of coupling, defined in article 124. Evidently α_1 and α_2 are the damping constants of the two circuits taken separately, and ω_1 and ω_2 are the angular resonance frequencies (90–5) of impressed e.m.f.'s for the individual circuits. For small damping, however, the resonance frequencies do not differ appreciably from the natural frequencies of free oscillation, as indicated by (89–22), so we can regard ω_1 and ω_2 as the independent angular frequencies of free oscillation.

We may expect damped harmonic oscillations as in the case

of the single oscillating circuit, so we look for solutions of the form $Ce^{\gamma t}$ where C is an arbitrary constant and $\gamma = -\alpha \pm i\omega_0$. Substituting $e^{(-\alpha+i\omega_0)t}$ for the current in either equation of (125–4) gives

$$\begin{aligned}(1-k^2)(-\alpha+i\omega_0)^4 &+ 2(\alpha_1+\alpha_2)(-\alpha+i\omega_0)^3 \\ &+ (\omega_1^2+4\alpha_1\alpha_2+\omega_2^2)(-\alpha+i\omega_0)^2 \\ &\quad + 2(\alpha_1\omega_2^2+\alpha_2\omega_1^2)(-\alpha+i\omega_0)+\omega_1^2\omega_2^2 = 0\end{aligned}$$

after the common factor $e^{(-\alpha+i\omega_0)t}$ has been divided out, and this equation reduces to

$$\begin{aligned}(1-k^2)(4i\alpha\omega_0^3+\omega_0^4) &+ 2(\alpha_1+\alpha_2)(-i\omega_0^3) \\ &+ (\omega_1^2+\omega_2^2)(-2i\alpha\omega_0-\omega_0^2)+2(\alpha_1\omega_2^2+\alpha_2\omega_1^2)(i\omega_0) \\ &\qquad + \omega_1^2\omega_2^2 = 0,\end{aligned}$$

when second and higher order terms in α are dropped. Separating real and imaginary parts, we have finally

$$(1-k^2)\omega_0^4-(\omega_1^2+\omega_2^2)\omega_0^2+\omega_1^2\omega_2^2 = 0, \qquad (125\text{–}5)$$

and

$$\begin{aligned}\{2(1-k^2)\omega_0^2-(\omega_1^2+\omega_2^2)\}\alpha &- \{(\alpha_1+\alpha_2)\omega_0^2 \\ &- (\alpha_1\omega_2^2+\alpha_2\omega_1^2)\} = 0. \qquad (125\text{–}6)\end{aligned}$$

Exactly the same equations are obtained, of course, if we substitute $e^{(-\alpha-i\omega_0)t}$ originally.

As there are two real positive roots of (125–5), namely,

$$\left.\begin{aligned} &= \sqrt{\frac{(\omega_1^2+\omega_2^2)-\sqrt{(\omega_1^2+\omega_2^2)^2-4(1-}}{2(1-k^2)}} \\ \omega_0 &= \sqrt{\frac{+\,\omega_2^2)+\sqrt{(\omega_1^2+\omega_2^2)^2-4(1-}}{2(1-k^2)}}\end{aligned}\right\}(125$$

there are two oscillation frequencies, both of which appear simultaneously in each circuit. From the form of (125–7) it is evident that these frequencies are always different from the characteristic frequencies ω_1 and ω_2 of the circuits oscillating independently. A special case of interest is when the circuits are *tuned* together,

so that $\omega_1 = \omega_2 \equiv \Omega$. Then (125–7) reduces to

$$\left.\begin{aligned} \omega_0' &= \frac{\Omega}{\sqrt{1+k}}, \\ \omega_0'' &= \frac{\Omega}{\sqrt{1-k}}. \end{aligned}\right\} \quad (125\text{–}8)$$

Corresponding to the two values of ω_0 there are two values of α, given by (125–6), which we may designate by α' and α'', respectively. These quantities have a simple form, independent of ω_1 and ω_2, only when the circuits are tuned together. In this case, using (125–8), we obtain from (125–6)

$$\left.\begin{aligned} \alpha' &= \frac{1}{1+k}\left(\frac{\alpha_1+\alpha_2}{2}\right), \\ \alpha'' &= \frac{1}{1-k}\left(\frac{\alpha_1+\alpha_2}{2}\right). \end{aligned}\right\} \quad (125\text{–}9)$$

Let us return now to the differential equations. The complete solutions are obtained by adding together the four possible expressions of the form $Ce^{\gamma t}$. Thus,

$$i_1 = A_1'e^{(-\alpha'+i\omega_0')t} + B_1'e^{(-\alpha'-i\omega_0')t} + A_1''e^{(-\alpha''+i\omega_0'')t} + B_1''e^{(-\alpha''-i\omega_0'')t},$$
$$i_2 = A_2'e^{(-\alpha'+i\omega_0')t} + B_2'e^{(-\alpha'-i\omega_0')t} + A_2''e^{(-\alpha''+i\omega_0'')t} + B_2''e^{(-\alpha''-i\omega_0'')t},$$

where the A's and B's are arbitrary constants. As in article 89 these equations may be put in trigonometric form and we find

$$\left.\begin{aligned} i_1 &= i_{10}'e^{-\alpha' t}\sin(\omega_0' t + \epsilon_1') + i_{10}''e^{-\alpha'' t}\sin(\omega_0'' t + \epsilon_1''), \\ i_2 &= i_{20}'e^{-\alpha' t}\sin(\omega_0' t + \epsilon_2') + i_{20}''e^{-\alpha'' t}\sin(\omega_0'' t + \epsilon_2''), \end{aligned}\right\} \quad (125\text{–}10)$$

the arbitrary constants now appearing as amplitudes and phase angles, represented by the i_0's and the ϵ's, respectively.

The arbitrary constants are not all independent, as may be seen by substituting (125–10) in either of the original differential equations. Thus, dividing the first equation of (125–2) by L_1 in order to introduce α_1, ω_1 and k, and substituting the values of

the currents from (125–10) gives, if we assume positive coupling,

$$
\begin{aligned}
(-\omega_0'^2 + \omega_1^2)i_{10}'e^{-\alpha' t} \sin(\omega_0' t + \epsilon_1')& \\
+ 2(-\alpha' + \alpha_1)\omega_0' i_{10}' e^{-\alpha' t} \cos(\omega_0' t + \epsilon_1')& \\
+ (-\omega_0''^2 + \omega_1^2) i_{10}'' e^{-\alpha'' t} \sin(\omega_0'' t + \epsilon_1'')& \\
+ 2(-\alpha'' + \alpha_1)\omega_0'' i_{10}'' e^{-\alpha'' t} \cos(\omega_0'' t + \epsilon_1'')& \\
= k\sqrt{\frac{L_2}{L_1}}\,\omega_0'^2 i_{20}' e^{-\alpha' t} \sin(\omega_0' t + \epsilon_2')& \\
+ 2k\sqrt{\frac{L_2}{L_1}}\,\alpha'\omega_0' i_{20}' e^{-\alpha' t} \cos(\omega_0' t + \epsilon_2')& \\
+ k\sqrt{\frac{L_2}{L_1}}\,\omega_0''^2 i_{20}'' e^{-\alpha'' t} \sin(\omega_0'' t + \epsilon_2'')& \\
+ 2k\sqrt{\frac{L_2}{L_1}}\,\alpha''\omega_0'' i_{20}'' e^{-\alpha'' t} \cos(\omega_0'' t + \epsilon_2''),&
\end{aligned}
$$

where the second order terms have been omitted as before. The equation holds for all values of t, so the parts of the equation in which ω_0' appears must represent one identity and the ω_0'' parts must represent another. Then, as the terms in α are negligible compared to those in ω,

$$
(-\omega_0'^2 + \omega_1^2)i_{10}'e^{-\alpha' t} \sin(\omega_0' t + \epsilon_1') = k\sqrt{\frac{L_2}{L_1}}\,\omega_0'^2 i_{20}' e^{-\alpha' t} \sin(\omega_0' t + \epsilon_2'),
$$

$$
(-\omega_0''^2 + \omega_1^2)i_{10}''e^{-\alpha'' t} \sin(\omega_0'' t + \epsilon_1'') = k\sqrt{\frac{L_2}{L_1}}\,\omega_0''^2 i_{20}'' e^{-\alpha'' t} \sin(\omega_0'' t + \epsilon_2''),
$$

from which we have

$$
(-\omega_0'^2 + \omega_1^2)i_{10}' = k\sqrt{\frac{L_2}{L_1}}\,\omega_0'^2 i_{20}', \qquad \epsilon_1' = \epsilon_2';
$$

$$
(-\omega_0''^2 + \omega_1^2)i_{10}'' = k\sqrt{\frac{L_2}{L_1}}\,\omega_0''^2 i_{20}'', \qquad \epsilon_1'' = \epsilon_2''.
$$

Hence, dropping unnecessary subscripts, the complete solutions

of (125–4) become

sin $\qquad\qquad$ sin $(\omega_0''t$

$$i_2 = \frac{1}{k}\sqrt{\frac{L_1}{L_2}}\left[\left(\frac{\omega_1}{\omega_0'}\right)^2 - 1\right] i_0' e^{-\alpha' t} \sin(\omega_0' t + \epsilon')$$

$$\sqrt{\frac{L_1}{L_2}}\Bigg[$$

The arbitrary constants are now independent and are determined by the initial conditions under which the oscillations are set up.

The general character of the oscillations is apparent from (125–11). There are two damped harmonic oscillations in each circuit the simultaneous occurrence of which leads to a periodic rise and fall of amplitude superposed on a steady falling off due to damping, in a manner analogous to the production of acoustic beats between two musical notes. Since $\omega_0' < \omega_1$, while $\omega_0'' > \omega_1$, the low frequency oscillations are in phase in the two circuits and the high frequency oscillations are π radians out of phase, so the electrical beats in the circuits are out of phase, as illustrated in Fig. 272. Evidently energy surges back and forth from one circuit to the other at the *beat frequency*, which is $(\omega_0'' - \omega_0')/2\pi$.

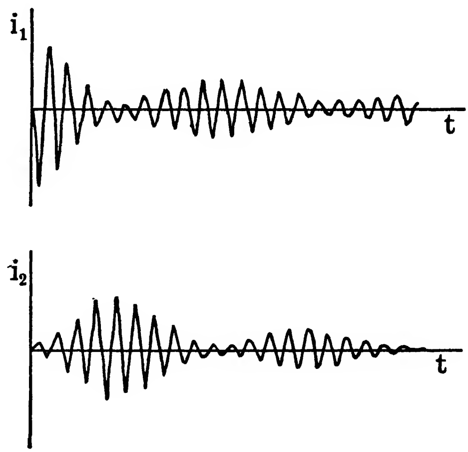


FIG. 272

The graphs in the figure represent oscillations produced by discharging the primary condenser at time $t = 0$. The values of the arbitrary constants for this case are found by setting

$$i_1 = 0, \qquad i_2 = 0,$$
$$q_1 = q_0, \qquad q_2 = 0,$$

for $t = 0$ in (125–11) and (125–1). The current conditions give

$$0 = i_0' \sin \epsilon' + i_0'' \sin \epsilon'',$$

$$0 = \frac{1}{k}\sqrt{\frac{L_1}{L_2}}\left[\left(\frac{\omega_1}{\omega_0'}\right)^2 - 1\right] i_0' \sin \epsilon' - \frac{1}{k}\sqrt{\frac{L_1}{L_2}}\left[1 - \left(\frac{\omega_1}{\omega_0''}\right)^2\right] i_0'' \sin \epsilon'',$$

from which we see that

$$\epsilon' = 0, \qquad \epsilon'' = 0. \tag{125–12}$$

The charge conditions lead to

$$\omega_0' i_0' \cos \epsilon' + \omega_0'' i_0'' \cos \epsilon'' + \omega_1^2 q_0 = -\left[\left(\frac{\omega_1}{\omega_0'}\right)^2 - 1\right]\omega_0' i_0' \cos \epsilon' + \left[1 - \left(\frac{\omega_1}{\omega_0''}\right)^2\right]\omega_0'' i_0'' \cos \epsilon'',$$

$$\left[\left(\frac{\omega_1}{\omega_0'}\right)^2 - 1\right]\omega_0' i_0' \cos \epsilon' - \left[1 - \left(\frac{\omega_1}{\omega_0''}\right)^2\right]\omega_0'' i_0'' \cos \epsilon'' = - k^2\omega_0' i_0' \cos \epsilon' - k^2 \omega_0'' i_0'' \cos \epsilon'',$$

after terms in α have been dropped. Using (125–12) and combining, these equations reduce to

$$\frac{1}{\omega_0'} i_0' + \frac{1}{\omega_0''} i_0'' + q_0 = 0,$$

$$\omega_0' i_0' + \omega_0'' i_0'' + \frac{\omega_1^2}{1 - k^2} q_0 = 0,$$

so that

$$i_0'(\omega_0''^2 - \omega_0'^2) = -\left(\omega_0''^2 - \frac{\omega_1^2}{1 - k^2}\right)\omega_0' q_0,$$

$$i_0''(\omega_0''^2 - \omega_0'^2) = -\left(\frac{\omega_1^2}{1 - k^2} - \omega_0'^2\right)\omega_0'' q_0.$$

Now, if we eliminate ω_2^2 between the equations (125–7), we have the relation

$$\frac{\omega_1^2}{1 - k^2} = \omega_0'^2 + \omega_0''^2 - \frac{\omega_0'^2 \omega_0''^2}{\omega_1^2},$$

which allows us to put the constants in the final form

$$\left.\begin{aligned} i_0' &= -\frac{1-\left(\dfrac{\omega_1}{\omega_0''}\right)^2}{\left(\dfrac{\omega_1}{\omega_0'}\right)^2-\left(\dfrac{\omega_1}{\omega_0''}\right)^2}\,\omega_0' q_0, \\ i_0'' &= -\frac{\left(\dfrac{\omega_1}{\omega_0'}\right)^2-1}{\left(\dfrac{\omega_1}{\omega_0'}\right)^2-\left(\dfrac{\omega_1}{\omega_0''}\right)^2}\,\omega_0'' q_0. \end{aligned}\right\} \quad (125\text{–}13)$$

When the circuits are tuned together the complete solutions corresponding to the initial conditions given above become

$$i_1 = -\frac{1}{2}\frac{\Omega}{\sqrt{1+k}}\,q_0 e^{-\alpha' t}\sin\frac{\Omega}{\sqrt{1+k}}t - \frac{1}{2}\frac{\Omega}{\sqrt{1-k}}\,q_0 e^{-\alpha'' t}\sin\frac{\Omega}{\sqrt{1-k}}t,$$

$$\begin{aligned} i_2 = -\frac{1}{2}\sqrt{\frac{L_1}{L_2}}\frac{\Omega}{\sqrt{1+k}}\,q_0 e^{-\alpha' t}\sin\frac{\Omega}{\sqrt{1+k}}t & \\ + \frac{1}{2}\sqrt{\frac{L_1}{L_2}}\frac{\Omega}{\sqrt{1-k}}\,q_0 e^{-\alpha'' t}\sin\frac{\Omega}{\sqrt{1-k}}t, & \end{aligned}$$

the damping constants being given by (125–9) above. Tuned coupled circuits of this sort find application in damped-wave wireless telegraphy. The circuit arrangements are shown in

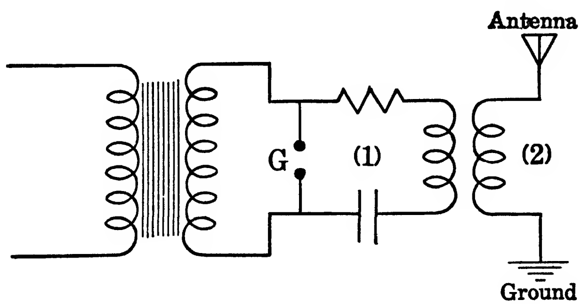


FIG. 273

Fig. 273. A spark gap G, connected to the secondary of a high voltage transformer, is included in the primary circuit (1). As the transformer voltage increases in magnitude the primary

condenser charges until finally a spark passes, effectively short-circuiting the gap, and oscillations ensue. As the frequencies are very high the oscillations die away long before the transformer voltage has passed through a half cycle. When the transformer voltage has reversed direction and again reaches sufficient magnitude another set of oscillations results, and so on, the number ν of wave trains generated per second being equal to twice the frequency of the transformer power supply. In the secondary circuit (2) the capacity is supplied by an antenna and ground system, from which energy is *radiated* in the form of electromagnetic waves, described in Chapter XVI.

The Wave-Meter. — In adjusting and using oscillating circuits it is often necessary to measure frequencies. This is most readily done by coupling to the source of oscillations whose frequency is to be determined an auxiliary adjustable circuit composed of an inductance, a variable condenser calibrated to give the frequency of the circuit, and an ammeter, usually of the thermocouple type, to show the mean-square current. Such an adjustable circuit, assembled as a unit, is called a *wave-meter*. The variable condenser is adjusted until the mean-square current is a maximum. Then, *provided the coupling is so small that the reaction of the wave-meter on the source of oscillations is negligible*, the frequency of the oscillations is equal to the frequency of the meter circuit, shown by the calibration of the condenser. It is necessary, of course, that the wave trains whose frequency is being determined should be repeated continuously, as in the case of the wireless signals described above, in order to obtain a steady deflection on the current meter.

To deduce the frequency relation stated above we require solutions of (125–4) valid in the vicinity of resonance for small coupling as well as small damping. The solutions (125–11) cannot be used as some of the terms are now of the same order as quantities neglected in the derivation of these solutions. We must therefore return to (125–4) and find new solutions, taking $k\omega$ and also $\omega_2 - \omega_1$ to be quantities at least of the first order of smallness.

As before, we know that the solution of (125–4) must be of the form $Ce^{\gamma t}$ where C is an arbitrary constant. Substituting this expression in the equation we find

$$\left\{\left(\gamma + \frac{\alpha_1 + \alpha_2}{2}\right)^2 + \frac{\omega_1{}^2 + \omega_2{}^2}{2} - \left(\frac{\alpha_1 + \alpha_2}{2}\right)^2\right\}^2$$
$$= \left\{(\alpha_2 - \alpha_1)\gamma + \frac{\omega_2{}^2 - \omega_1{}^2}{2}\right\}^2 + k^2\gamma^4 \qquad (125\text{–}14)$$

without neglect of any terms. Putting

$$A \equiv \frac{\alpha_1 + \alpha_2}{2}, \quad \Omega^2 \equiv \frac{\omega_1{}^2 + \omega_2{}^2}{2} - A^2,$$

$$\Delta\alpha \equiv \alpha_2 - \alpha_1, \quad \Delta\omega \equiv \frac{\omega_2{}^2 - \omega_1{}^2}{2\Omega},$$

this simplifies to

$$\{(\gamma + A)^2 + \Omega^2\}^2 = \{\Omega\Delta\omega + \gamma\Delta\alpha\}^2 + k^2\gamma^4. \qquad (125\text{–}15)$$

We must solve this equation for the case where α_1 and α_2 are small quantities of the first order as compared with Ω, while $\Delta\alpha$, $\Delta\omega$ and $k\Omega$ are of the same order of magnitude as α_1 and α_2 or smaller. Note that under these conditions $\Delta\omega = \omega_2 - \omega_1$ to a high degree of accuracy.

To get a first approximation to the solution we neglect all the small quantities, obtaining $\gamma^2 + \Omega^2 = 0$. This gives $\gamma = i\Omega$, where we have retained only the positive root since we are interested only in the magnitudes of the frequencies. As each term in γ on the right of (125–15) contains the square or product of two of the small quantities $\Delta\omega$, $\Delta\alpha$, k, we can substitute this approximate value there and solve the equation to obtain a second approximation. We have then

$$\{(\gamma + A)^2 + \Omega^2\}^2 = \Omega^2\{(\Delta\omega)^2 - (\Delta\alpha)^2 + k^2\Omega^2 + 2i\Delta\omega\Delta\alpha\}$$
$$= \Omega^2 p^2 e^{2i\psi} \qquad (125\text{–}16)$$

where

$$p^2 \equiv \sqrt{\{(\Delta\omega)^2 - (\Delta\alpha)^2 + k^2\Omega^2\}^2 + 4(\Delta\omega)^2(\Delta\alpha)^2},$$

$$\tan 2\psi \equiv \frac{2\Delta\omega\Delta\alpha}{(\Delta\omega)^2 - (\Delta\alpha)^2 + k^2\Omega^2}.$$

Taking the square root of both sides,

$$(\gamma + A)^2 = -\,\Omega^2\left(1 \mp \frac{p}{\Omega}e^{i\psi}\right),$$

and extracting the positive root as we are concerned only with the magnitudes of the frequencies,

$$\gamma = -A + i\Omega\left(1 \mp \frac{p}{2\Omega} e^{i\psi}\right) \tag{125-17}$$

since p is small compared to Ω. Separating the real and imaginary parts for each value of the double sign we find the following damping constants (α', α'') and corresponding frequencies (ω_0', ω_0''):

$$\left.\begin{aligned} \alpha' &= A - \tfrac{1}{2}p \sin\psi, \quad & \omega_0' &= \Omega - \tfrac{1}{2}p \cos\psi, \\ \alpha'' &= A + \tfrac{1}{2}p \sin\psi, \quad & \omega_0'' &= \Omega + \tfrac{1}{2}p \cos\psi. \end{aligned}\right\} \tag{125-18}$$

Evidently $\alpha' + \alpha'' = 2A = \alpha_1 + \alpha_2$. Both frequencies with their damping constants appear in each circuit, as in the case previously analyzed, and the currents have exactly the same general trigonometric form, namely,

$$\left.\begin{aligned} i_1 &= i_{10}'e^{-\alpha' t} \sin(\omega_0' t + \epsilon_1') + i_{10}''e^{-\alpha'' t} \sin(\omega_0'' t + \epsilon_1''), \\ i_2 &= i_{20}'e^{-\alpha' t} \sin(\omega_0' t + \epsilon_2') + i_{20}''e^{-\alpha'' t} \sin(\omega_0'' t + \epsilon_2''). \end{aligned}\right\} \tag{125-19}$$

The four relations between the eight arbitrary constants are found exactly as before (p. 507), the condition equations in this case reducing to

$$\begin{aligned} (-\omega_0'^2 + \omega_1^2)i_{10}'e^{-\alpha' t} &\sin(\omega_0' t + \epsilon_1') \\ &+ 2(-\alpha' + \alpha_1)\omega_0' i_{10}'e^{-\alpha' t} \cos(\omega_0' t + \epsilon_1') \\ &\qquad = k\sqrt{\frac{L_2}{L_1}}\, \omega_0'^2 i_{20}'e^{-\alpha' t} \sin(\omega_0' t + \epsilon_2'), \end{aligned}$$

$$\begin{aligned} (-\omega_0''^2 + \omega_1^2)i_{10}''e^{-\alpha'' t} &\sin(\omega_0'' t + \epsilon_1'') \\ &+ 2(-\alpha'' + \alpha_1)\omega_0'' i_{10}''e^{-\alpha'' t} \cos(\omega_0'' t + \epsilon_1'') \\ &\qquad = k\sqrt{\frac{L_2}{L_1}}\, \omega_0''^2 i_{20}''e^{-\alpha'' t} \sin(\omega_0'' t + \epsilon_2''). \end{aligned}$$

Consider the first condition equation. As it must hold for all values of the time, we must have

$$(-\omega_0'^2 + \omega_1^2)i_{10}' = k\sqrt{\frac{L_2}{L_1}}\, \omega_0'^2 i_{20}' \cos(\epsilon_2' - \epsilon_1'),$$

$$2(-\alpha' + \alpha_1)\omega_0' i_{10}' = k\sqrt{\frac{L_2}{L_1}}\, \omega_0'^2 i_{20}' \sin(\epsilon_2' - \epsilon_1').$$

Now, to the required degree of approximation,

$$-\omega_0'^2 + \omega_1^2 = \Omega(p\cos\psi - \Delta\omega),$$

$$2(-\alpha' + \alpha_1)\omega_0' = \Omega(p\sin\psi - \Delta\alpha),$$

giving

$$\frac{i_{10}'}{i_{20}'} = \frac{k\Omega}{r'}\sqrt{\frac{L_2}{L_1}}, \quad r' \equiv \sqrt{(p\sin\psi - \Delta\alpha)^2 + (p\cos\psi - \Delta\omega)^2},$$

and

$$\sin(\epsilon_2' - \epsilon_1') = \frac{p\sin\psi - \Delta\alpha}{r'}, \quad \cos(\epsilon_2' - \epsilon_1') = \frac{p\cos\psi - \Delta\omega}{r'}.$$

Similarly the second condition equation gives

$$\frac{i_{10}''}{i_{20}''} = \frac{k\Omega}{r''}\sqrt{\frac{L_2}{L_1}}, \quad r'' \equiv \sqrt{(p\sin\psi + \Delta\alpha)^2 + (p\cos\psi + \Delta\omega)^2},$$

and

$$\sin(\epsilon_2'' - \epsilon_1'') = -\frac{p\sin\psi + \Delta\alpha}{r''},$$

$$\cos(\epsilon_2'' - \epsilon_1'') = -\frac{p\cos\psi + \Delta\omega}{r''}.$$

After some reduction we find that $r'r'' = k^2\Omega^2$. For the moment let us suppose that $\Delta\alpha$ and $\Delta\omega$ are both positive, so that $r'' > r'$. Then we see r'' is of the order of magnitude of $\sqrt{(\Delta\alpha)^2 + (\Delta\omega)^2 + k^2\Omega^2}$ and r' is of the order of magnitude of $k^2\Omega^2/\sqrt{(\Delta\alpha)^2 + (\Delta\omega)^2 + k^2\Omega^2}$. So, although i_{10}' may be large compared with i_{20}', nevertheless $k\sqrt{L_1/L_2}\, i_{10}'$ is always small compared with i_{20}'. Similarly $k\sqrt{L_1/L_2}\, i_{10}''$ is always small compared with i_{20}''. If either or both $\Delta\alpha$ and $\Delta\omega$ are negative, similar reasoning leads to the same conclusions.

It remains to determine the arbitrary constants corresponding to given initial conditions and then to calculate the mean-square secondary current. In order to deal with any type of excitation let us suppose that initially in the primary the current is I_1 and its rate of change is T_1 while in the secondary the current is zero and the charge on the condenser is zero, that is, referring to (125–1),

the rate of change of current is $-k\sqrt{L_1/L_2}T_1$. Thus, putting

$$i_1 = I_1, \qquad \frac{di_1}{dt} = T_1,$$

$$i_2 = 0, \qquad \frac{di_2}{dt} = -k\sqrt{\frac{L_1}{L_2}}\,T_1,$$

for $t = 0$, and making the usual approximations, we have

$$\left.\begin{aligned} i_{10}'\sin\epsilon_1' + i_{10}''\sin\epsilon_1'' &= I_1,\\ \omega_0'i_{10}'\cos\epsilon_1' + \omega_0''i_{10}''\cos\epsilon_1'' &= T_1,\\ i_{20}'\sin\epsilon_2' + i_{20}''\sin\epsilon_2'' &= 0,\\ \omega_0'i_{20}'\cos\epsilon_2' + \omega_0''i_{20}''\cos\epsilon_2'' &= -k\sqrt{\frac{L_1}{L_2}}\,T_1.\end{aligned}\right\}\quad(125\text{–}21)$$

As we may replace the nearly equal frequencies ω_0' and ω_0'' by Ω, the last two equations give

$$i_{20}'^2 + 2i_{20}'i_{20}''\cos(\epsilon_2' - \epsilon_2'') + i_{20}''^2 = k^2\left(\frac{L_1}{L_2}\right)\frac{T_1^2}{\Omega^2}.$$

The right-hand side of this equation has the order of magnitude of the larger of $k^2(L_1/L_2)i_{10}'^2$ and $k^2(L_1/L_2)i_{10}''^2$. As these are negligible compared with individual terms on the other side of the equation,

$$i_{20}'^2 + 2i_{20}'i_{20}''\cos(\epsilon_2' - \epsilon_2'') + i_{20}''^2 = 0,$$

or

$$(i_{20}' - i_{20}'')^2 = -4i_{20}'i_{20}''\cos^2\left(\frac{\epsilon_2' - \epsilon_2''}{2}\right).$$

We shall consider all current amplitudes to be essentially positive, allowing phase differences to assume any required values between 0 and 2π. Then the preceding equation can be satisfied only by

$$i_{20}' = i_{20}'' \equiv i_{20}, \quad \epsilon_2' = \epsilon_2'' + \pi \equiv \epsilon_2.$$

Now, from the first two relations in (125–21),

$$i_{10}'^2 + 2i_{10}'i_{10}''\cos(\epsilon_1' - \epsilon_1'') + i_{10}''^2 = I_1^2 + \frac{T_1^2}{\Omega^2}.$$

But

$$\begin{aligned}\cos(\epsilon_1' - \epsilon_1'') &= -\cos\{(\epsilon_2'' - \epsilon_1'') - (\epsilon_2' - \epsilon_1')\}\\ &= \frac{\{p^2\cos^2\psi - (\Delta\alpha)^2\} + \{p^2\sin^2\psi - (\Delta\alpha)^2\}}{r'r''}.\end{aligned}$$

Hence, making use of the relations between i_{10}', i_{10}'' and i_{20}', i_{20}'' previously deduced,

$$\frac{i_{20}^2}{k^2\Omega^2}\left(\frac{L_2}{L_1}\right)[r''^2 + 2\{p^2\cos^2\psi - (\Delta\omega)^2\} + 2\{p^2\sin^2\psi - (\Delta\alpha)^2\} + r'^2]$$
$$= I_1^2 + \frac{T_1^2}{\Omega^2}.$$

Referring to the expressions for r' and r'', we see that the factor in the brackets reduces to $4p^2 = 4\{(\alpha'' - \alpha')^2 + (\omega_0'' - \omega_0')^2\}$.

So

$$i_{20}^2 = \frac{k^2\Omega^2}{4}\left(\frac{L_1}{L_2}\right)\frac{I_1^2 + \dfrac{T_1^2}{\Omega^2}}{(\alpha'' - \alpha')^2 + (\omega_0'' - \omega_0')^2}. \quad (125\text{–}23)$$

Let us now calculate the mean-square secondary current, assuming there are ν wave-trains per second. Expressing the values of the constants in terms of i_{20} and ϵ_2 we have

$$\overline{i_2^2} = \nu\int_0^\infty i_{20}^2\,[\tfrac{1}{2}e^{-2\alpha' t}\{1 - \cos 2(\omega_0' t + \epsilon_2)\}$$
$$- e^{-(\alpha'+\alpha'')t}\{\cos(\omega_0'' - \omega_0')t - \cos[(\omega_0' + \omega_0'')t + 2\epsilon_2]\}$$
$$+ \tfrac{1}{2}e^{-2\alpha'' t}\{1 - \cos 2(\omega_0'' t + \epsilon_2)\}]\,dt.$$

From the integration formula,

$$\int_0^\infty e^{-\alpha t}\cos(\omega t + \epsilon)dt = \frac{\alpha\cos\epsilon + \omega\sin\epsilon}{\alpha^2 + \omega^2},$$

we see that the large frequency terms may be neglected and that

$$\overline{i_2^2} = \nu i_{20}^2\left[\frac{1}{4\alpha'} - \frac{\alpha'' + \alpha'}{(\alpha'' + \alpha')^2 + (\omega_0'' - \omega_0')^2} + \frac{1}{4\alpha''}\right]$$
$$= \nu i_{20}^2\,\frac{\alpha' + \alpha''}{\alpha'\alpha''}\;\frac{(\alpha'' - \alpha')^2 + (\omega_0'' - \omega_0')^2}{(\alpha'' + \alpha')^2 + (\omega_0'' - \omega_0')^2}.$$

Now

$$\alpha'\alpha''\{(\alpha'' + \alpha')^2 + (\omega_0'' - \omega_0')^2\}$$
$$= 4(A^2 - \tfrac{1}{4}p^2\sin^2\psi)(A^2 + \tfrac{1}{4}p^2\cos^2\psi)$$
$$= 4\{A^4 + \tfrac{1}{4}p^2A^2\cos 2\psi - \tfrac{1}{64}p^4\sin^2 2\psi\}$$
$$= \alpha_1\alpha_2\{(\alpha_1 + \alpha_2)^2\left(1 + \frac{k^2\Omega^2}{4\alpha_1\alpha_2}\right) + (\Delta\omega)^2\}.$$

Substituting the value of i_{20}^2 and remembering that $\alpha' + \alpha'' = \alpha_1 + \alpha_2$, we have finally

$$\overline{i_2^2} = \frac{\nu k^2\left(\frac{L_1}{L_2}\right)(I_1^2\Omega^2 + T_1^2)}{16\alpha_1\alpha_2}\left[\frac{\alpha_1 + \alpha_2}{(\alpha_1 + \alpha_2)^2\left(1 + \frac{k^2\Omega^2}{4\alpha_1\alpha_2}\right) + (\Delta\omega)^2}\right],$$

or in the limiting case where $k\Omega$ is small compared with α_1 and α_2,

$$\overline{i_2^2} = \frac{\nu k^2\left(\frac{L_1}{L_2}\right)(I_1^2\Omega^2 + T_1^2)}{16\alpha_1\alpha_2}\left[\frac{\alpha_1 + \alpha_2}{(\alpha_1 + \alpha_2)^2 + (\Delta\omega)^2}\right]. \quad (125\text{–}24)$$

When the oscillations are produced by discharging the primary condenser, as described above (p. 511), $I_1 = 0$ and $T_1 = -q_0\Omega^2$. A more common method of excitation for measurement purposes, where small power is involved, is to discharge the primary inductance. That is, by means of a vibrator V (Fig. 274) a steady current $I_1 = -\mathcal{E}_0/R_1$ is established through

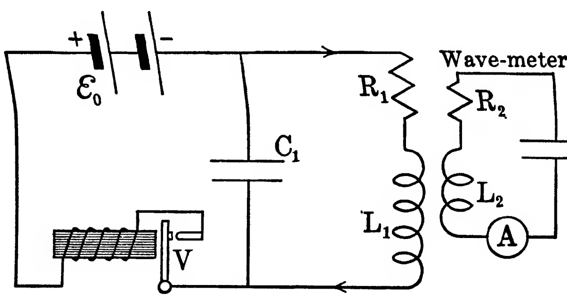


FIG. 274

R_1 and L_1, then suddenly interrupted. Here I_1 has the value just given and $T_1 = 0$. With any type of excitation, however, a maximum value of $\overline{i_2^2}$ corresponds to $\Delta\omega = 0$, that is, to $\omega_1 = \omega_2$, which is the necessary condition for use of the wave-meter.

Problem 125a. Given an oscillogram of the current in either of two coupled circuits tuned together to the angular frequency Ω and coupled loosely enough for k^2 to be negligible, find a simple expression

for k in terms of quantities which may be obtained from the oscillogram. Ans. $k = n/N$, where n is the beat frequency and N is the mean oscillation frequency.

Problem 125b. Taking ω_2 constant, plot ω_0' and ω_0'' as a function of ω_1 for k equal to 0.1, 0.5 and 0.9. Compare the three sets of curves.

126. Inductively Coupled Circuits in Forced Oscillation. — Let us now investigate the behavior of coupled circuits (Fig. 275) when an alternating e.m.f. $\mathcal{E} = \mathcal{E}_0 \sin \omega t$ is included in the primary. The circuit equations are

$$\left.\begin{aligned} L_1 \frac{di_1}{dt} + R_1 i_1 + \frac{q_1}{C_1} &= -M\frac{di_2}{dt} + \mathcal{E}_0 \sin \omega t, \\ L_2 \frac{di_2}{dt} + R_2 i_2 + \frac{q_2}{C_2} &= -M\frac{di_1}{dt}. \end{aligned}\right\} \quad (126\text{–}1)$$

Assuming that the e.m.f. has been acting long enough for a steady state to have been reached, the currents are sinusoidal in

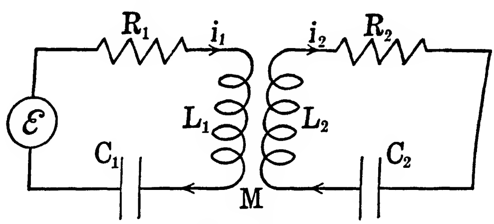


FIG. 275

form with an angular frequency ω. Hence we may conveniently use the complex method of analysis, described in article 112, to obtain the solutions of (126–1). Replacing $\mathcal{E}$ by $\boldsymbol{\mathcal{E}} = \mathcal{E}_0 e^{\iota\omega t}$ and introducing complex currents we have, after differentiation,

$$\left.\begin{aligned} L_1 \frac{d^2\mathbf{i}_1}{dt^2} + R_1 \frac{d\mathbf{i}_1}{dt} + \frac{1}{C_1}\mathbf{i}_1 &= -M\frac{d^2\mathbf{i}_2}{dt^2} + \frac{d}{dt}(\mathcal{E}_0 e^{\iota\omega t}), \\ L_2 \frac{d^2\mathbf{i}_2}{dt^2} + R_2 \frac{d\mathbf{i}_2}{dt} + \frac{1}{C_2}\mathbf{i}_2 &= -M\frac{d^2\mathbf{i}_1}{dt^2}, \end{aligned}\right\} \quad (126\text{–}2)$$

corresponding to (112–1) for a single circuit.

The complex currents must have the form

$$\mathbf{i}_1 = \mathbf{i}_{10} e^{\iota\omega t}, \qquad \mathbf{i}_2 =$$

Substituting these in (126–2) and removing common factors we have

$$\left.\begin{aligned}\left[R_1 + \mathrm{i}\left(L_1\omega - \frac{1}{C_1\omega}\right)\right]\mathbf{i}_{10} + \mathrm{i}M\omega\mathbf{i}_{20} &= \mathcal{E}_0,\\ \left[R_2 + \mathrm{i}\left(L_2\omega - \frac{1}{C_2\omega}\right)\right]\mathbf{i}_{20} + \mathrm{i}M\omega\mathbf{i}_{10} &= 0,\end{aligned}\right\} \qquad (126\text{–}3)$$

which give

$$\left.\begin{aligned}\mathbf{i}_{10} &= \frac{\mathcal{E}_0}{(R_1 + \mathrm{i}X_1) + \dfrac{M^2\omega^2}{R_2 + \mathrm{i}X_2}} = \frac{\mathcal{E}_0}{\mathbf{Z}_1 + \dfrac{M^2\omega^2}{\mathbf{Z}_2}},\\ \mathbf{i}_{20} &= \frac{-\mathrm{i}M\omega\mathcal{E}_0}{(R_2 + \mathrm{i}X_2)\left[(R_1 + \mathrm{i}X_1) + \dfrac{M^2\omega^2}{R_2 + \mathrm{i}X_2}\right]}\\ &= \frac{-\mathrm{i}M\omega\mathcal{E}_0}{\mathbf{Z}_2\left[\mathbf{Z}_1 + \dfrac{M^2\omega^2}{\mathbf{Z}_2}\right]},\end{aligned}\right\} \qquad (126\text{–}4)$$

where X_1, X_2 are the reactances of the separate circuits, and $\mathbf{Z}_1$, $\mathbf{Z}_2$ are the corresponding complex impedances. The same result may be obtained by applying Kirchhoff's laws to the coupled circuits regarded as a network. In this case we have

$$\begin{aligned}\mathbf{Z}_1\mathbf{i}_1 + \mathrm{i}M\omega\mathbf{i}_2 &= \boldsymbol{\mathcal{E}},\\ \mathbf{Z}_2\mathbf{i}_2 + \mathrm{i}M\omega\mathbf{i}_1 &= 0,\end{aligned}$$

which are equivalent to (126–3).

In order to find the actual currents we must take the imaginary parts of $\mathbf{i}_1$ and $\mathbf{i}_2$. Setting

$$Z_1' e^{\mathrm{i}\phi_1'} \equiv \mathbf{Z}_1 + \frac{M^2\omega^2}{\mathbf{Z}_2},$$

equations (126–4) become

$$\mathbf{i}_{10} = \frac{\mathcal{E}_0}{Z_1'}e^{-\mathrm{i}\phi_1'},$$

$$\mathbf{i}_{20} = \frac{-\mathrm{i}M\omega\mathcal{E}_0}{Z_2 Z_1'}e^{-\mathrm{i}(\phi_2+\phi_1')} = \frac{M\omega\mathcal{E}_0}{Z_2 Z_1'}e^{-\mathrm{i}(\phi_2+\phi_1'+\pi/2)},$$

where

$$Z_1' = \sqrt{\left(R_1 + \frac{M^2\omega^2}{Z_2^2} R_2\right)^2 + \left(X_1 - \frac{M^2\omega^2}{Z_2^2} X_2\right)^2},$$

$$\tan \phi_1' = \frac{X_1 - \dfrac{M^2\omega^2}{Z_2^2} X_2}{R_1 + \dfrac{M^2\omega^2}{Z_2^2} R_2}. \quad (126\text{–}5)$$

Hence

$$\left.\begin{aligned} i_1 &= \frac{\mathcal{E}_0}{Z_1'} \sin (\omega t - \phi_1'), \\ i_2 &= \frac{M\omega\mathcal{E}_0}{Z_2 Z_1'} \sin \left(\omega t - \phi_2 - \phi_1' - \frac{\pi}{2}\right). \end{aligned}\right\} \quad (126\text{–}6)$$

We are concerned chiefly with the current amplitudes,

$$i_{10} = \frac{\mathcal{E}_0}{Z_1'}, \qquad i_{20} = \frac{M\omega\mathcal{E}_0}{Z_2 Z_1'}.$$

As these are functions both of the applied frequency and of the constants of the circuits, we shall investigate their behavior when ω alone is varied and also when X_1 and X_2 are changed with ω constant. In practice the resistances are almost always small compared to the reactances, except near the independent resonance points where the latter vanish. We shall therefore confine our analysis to the case of small resistances.

Frequency Variation. — Consider first i_{20} as it is the more important of the two amplitudes. We can write

$$Z_2 Z_1' = \sqrt{(R_1^2 + X_1^2)(R_2^2 + X_2^2) + 2M^2\omega^2(R_1R_2 - X_1X_2) + M^4\omega^4}. \quad (126\text{–}7)$$

Neglecting the resistances in comparison with the reactances, this vanishes when

$$X_1X_2 - M^2\omega^2 = 0,$$

that is, when

$$(1 - k^2)\omega^4 - (\omega_1^2 + \omega_2^2)\omega^2 + \omega_1^2\omega_2^2 = 0, \quad (126\text{–}8)$$

the constants k, ω_1 and ω_2 having the same significance as in

(125–5). Evidently equation (126–8) has the same roots as (125–5), since the equations are identical in form.

Now, maximum values of i_{20} correspond, very closely at least, to minimum values of Z_2Z_1'. Thus there are two *resonance frequencies* for the secondary circuit, given by

$$\left.\begin{aligned} \omega_r' &= \sqrt{\frac{(\omega_1^2+\omega_2^2)-\sqrt{(\omega_1^2+\omega_2^2)^2-4(1-k^2)\omega_1^2\omega_2^2}}{2(1-k^2)}}, \\ \omega_r'' &= \sqrt{\frac{(\omega_1^2+\omega_2^2)+\sqrt{(\omega_1^2+\omega_2^2)^2-4(1-k^2)\omega_1^2\omega_2^2}}{2(1-k^2)}}, \end{aligned}\right\} \quad (126\text{–}9)$$

and these are identical with the frequencies of free oscillation. When the circuits are tuned together, as is often the case, so that $\omega_1 = \omega_2 \equiv \Omega$, we have simply

$$\left.\begin{aligned} \omega_r' &= \frac{\Omega}{\sqrt{1+k}}, \\ \omega_r'' &= \frac{\Omega}{\sqrt{1-k}}. \end{aligned}\right\} \quad (126\text{–}10)$$

Maximum values of i_{10} occur at very nearly the same values of ω as maximum values of i_{20}. The factor $M\omega/Z_2$ is increasing at ω_r' and decreasing at ω_r'', so that the primary maxima are slightly farther apart than the secondary. The chief effect of

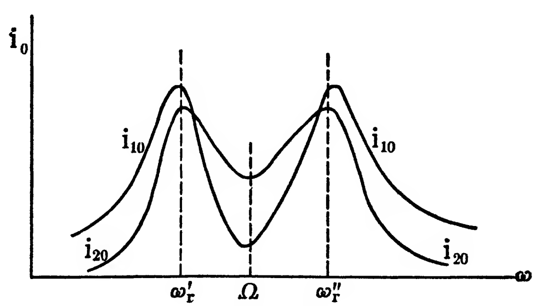


FIG. 276

$M\omega/Z_2$ is, however, to make i_{10} decrease more rapidly than i_{20} between the maxima and less rapidly outside. The forms of the *resonance curves* for two circuits tuned together are shown in

Fig. 276. In this case the magnitudes of the maxima are the same in each circuit. Note that, in accord with (126–10), the resonance peaks are not symmetrically placed relative to Ω.

If the coupling is made small the foregoing results are somewhat modified, for the resonance values of ω approach ω_1 and ω_2 so that the reactances are no longer large compared to the resistances. With the circuits tuned together the two resonance peaks merge into one as M diminishes, the coalition being more rapid in the secondary than in the primary because of the factor $M\omega/Z_2$. Resonance curves for this case are shown in Fig. 277.

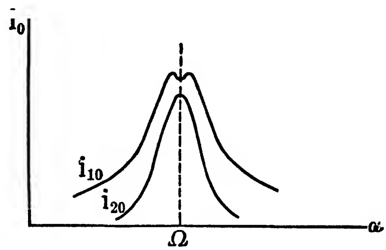


FIG. 277

The behavior of the secondary resonance at small coupling enables us to adjust the frequency of a variable source by means of a wave-meter (art. 125). The wave-meter, comprising the secondary circuit, is set for a desired angular frequency Ω and coupled loosely to the primary circuit containing the variable frequency ω. The primary is tuned until variation of ω produces only a single resonance peak in the secondary as indicated by the current-square meter of the wave-meter. Then the frequency adjustment giving maximum indication on the meter corresponds to $\omega = \Omega$.

Circuit Tuning. — When X_1 and X_2 are varied by means of the circuit elements, ω being kept constant, the exact conditions for resonance can be found without difficulty. Maximum values of i_{20} occur at minimum values of Z_2Z_1', which may be found by setting the partial derivatives with respect to X_1 and X_2 equal to zero. Thus,

$$\left.\begin{aligned} \frac{\partial}{\partial X_1}(Z_2Z_1') &= \frac{X_1(R_2^2 + X_2^2) - M^2\omega^2 X_2}{Z_2Z_1'} = 0, \\ \frac{\partial}{\partial X_2}(Z_2Z_1') &= \frac{(R_1^2 + X_1^2)X_2 - M^2\omega^2 X_1}{Z_2Z_1'} = 0, \end{aligned}\right\} \quad (126\text{–}11)$$

which are satisfied either by

$$X_1 = 0, \qquad X_2 = 0, \tag{126–12}$$

or by

$$\left. \begin{aligned} X_1 &= \pm\sqrt{\frac{R_1}{R_2}(M^2\omega^2 - R_1R_2)}, \\ X_2 &= \pm\sqrt{\frac{R_2}{R_1}(M^2\omega^2 - R_1R_2)}. \end{aligned} \right\} \tag{126–13}$$

In the latter case both positive signs must be used, or both negative, so there are two possible sets of values.

Applying the usual tests for minima we find that (126–12) represents the conditions for secondary resonance for the case $M^2\omega^2 < R_1R_2$, which we shall designate as *insufficient coupling.* The corresponding value of i_{20} is

$$(i_{20})_{\max} = \frac{M\omega\mathcal{E}_0}{M^2\omega^2 + R_1R_2}, \tag{126–14}$$

which vanishes if $M = 0$ as is to be expected.

On the other hand (126–13) gives the resonance conditions for $M^2\omega^2 > R_1R_2$, which we shall call *sufficient coupling.* In this case, for either the positive or the negative set of values,

$$(i_{20})_{\max} = \frac{\mathcal{E}_0}{2\sqrt{R_1R_2}}. \tag{126–15}$$

Thus with sufficient coupling the maximum current obtainable in the secondary is entirely independent of the coupling (Fig. 278) and, incidentally, of the frequency. Tuning the circuits for secondary resonance with the reactances positive is often referred to as *long-wave tuning* and with the reactances negative as *short-wave tuning.* The latter is usually more convenient as smaller variable condensers or inductances are required.

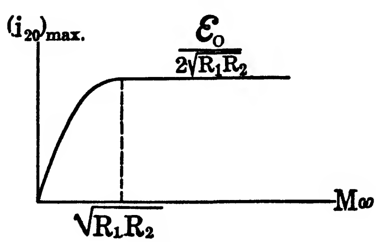


FIG. 278

As regards the primary circuit, there is no complete resonance, the maximum value of i_{10} increasing continuously as X_2 ap-

proaches infinity. There is, however, a partial resonance corresponding to adjustment of X_1 only, given by

$$\frac{\partial Z_1'}{\partial X_1} = \frac{X_1 - \dfrac{M^2\omega^2}{Z_2^2} X_2}{Z_1'} = 0.$$

This is equivalent to the first equation of (126–11), so the primary partial resonance for X_1 occurs at the same place as the secondary partial resonance for X_1.

The frequency ω of a sinusoidal e.m.f. may be determined with a wave-meter using the same procedure as in the case of free oscillations (art. 125). That is, the wave-meter is loosely coupled to the circuit containing the e.m.f. and adjusted until the current-square meter shows a maximum. Then we must have $\omega = \omega_2$, the frequency indicated by the wave-meter setting, for with M small

$$i_{20} = \frac{M\omega\mathcal{E}_0}{\sqrt{(R_1^2 + X_1^2)(R_2^2 + X_2^2)}} \qquad (126\text{–}16)$$

and the maximum occurs when $X_2 = 0$, that is, when the impressed angular frequency equals $1/\sqrt{L_2C_2}$. Tuning the primary increases the meter deflection for any given coupling, but is not necessary to establish the secondary resonance relation.

Problem 126a. Referring to Fig. 275, suppose the e.m.f. is connected across C_1 instead of in series with it. Neglecting the resistances R_1 and R_2, find the condition for resonance, that is, the condition under which the equivalent impedance of the network is a minimum.

$$\text{Ans. } L_1\omega\left(L_2\omega - \frac{1}{C_2\omega}\right) = M^2\omega^2.$$

Problem 126b. Deduce the transformer equations (117–2) from (126–4).

Problem 126c. Given two identical series circuits with resistance 10 ohms, inductance 20 millihenries and capacity variable. They are coupled inductively, the coefficient of coupling being 0.1. If a sinusoidal e.m.f. of frequency $(10)^4$ cycles is introduced into one of the circuits, find how the capacities must be adjusted to give a maximum current in the other circuit. Ans. $C_1 = C_2 = 0.0141$ microfarad, or $C_1 = C_2 = 0.0115$ microfarad.

127. Measurement of Logarithmic Decrement. — Knowledge of the logarithmic decrement δ of a circuit is often essential. From its definition (89–24) we see that it is an index of the rapidity with which free oscillations in the circuit are damped out. Again, in the case of forced oscillations it indicates the sharpness of resonance, provided, as usual, the damping is not large. For consider a series circuit which can be tuned by means of a variable capacity to resonance with an applied e.m.f. of fixed angular frequency ω. The ratio of the amplitude of the current when the circuit is not adjusted to resonance to the amplitude at resonance is given by

$$\frac{i_0}{(i_0)_{\max}} = \frac{R}{\sqrt{R^2 + \left(L\omega - \frac{1}{C\omega}\right)^2}} = \frac{1}{\sqrt{1 + \left(\frac{L\omega}{R}\right)^2 \left(1 - \frac{\omega_0{}^2}{\omega^2}\right)^2}},$$

where $\omega_0 \equiv 1/\sqrt{LC}$ (the variable in the equation) is the angular frequency of free oscillation corresponding to the given adjustment. Now the logarithmic decrement for the resonance adjustment $\omega_0 = \omega$ is $\delta = \pi R/L\omega$ from (89–25), so that

$$\frac{L\omega}{R} = \frac{\pi}{\delta}.$$

Thus,

$$\frac{i_0}{(i_0)_{\max}} = \frac{1}{\sqrt{1 + \left(\frac{\pi}{\delta}\right)^2 \left(1 - \frac{\omega_0{}^2}{\omega^2}\right)^2}},$$

which shows that the current falls off the more rapidly with varying ω_0 on either side of resonance the smaller the logarithmic decrement at resonance.

It is usually difficult to calculate logarithmic decrements accurately, particularly at high frequencies where the apparent values of the circuit elements may depend on the frequency, so we must resort to direct measurements. These may be made with either free oscillations or forced oscillations, that is, with damped waves or with undamped waves. Both methods will be discussed.

Measurements with Damped Waves. — Let us suppose that a circuit composed of an inductance, a variable condenser and a current-square meter, similar to a wave-meter circuit, is coupled loosely to a source of damped waves. Then, from (125–24), the mean-square current indicated by the meter is

$$\overline{i_2^2} = \frac{\nu k^2 \left(\frac{L_1}{L_2}\right)(I_1^2\Omega^2 + T_1^2)}{16\alpha_1\alpha_2}\left[\frac{\alpha_1 + \alpha_2}{(\alpha_1 + \alpha_2)^2 + (\Delta\omega)^2}\right],$$

where α_1 and α_2 are the damping constants associated with the source of oscillations and with the auxiliary measuring circuit, respectively, and $\Delta\omega = \omega_2 - \omega_1$ is the difference between the characteristic angular frequency of the measuring circuit at the given adjustment and the angular frequency of the source.

Now consider the mean-square current, designated by $\mathcal{J}$ for simplicity, as a function of the tuning, that is, as a function of ω_2. The maximum value $\mathcal{J}_m$ occurs when $\omega_2 = \omega_1 \equiv \omega_m$. Evidently

$$\frac{\mathcal{J}}{\mathcal{J}_m} = \frac{(\alpha_1 + \alpha_2)^2}{(\alpha_1 + \alpha_2)^2 + (\Delta\omega)^2}, \tag{127–1}$$

and the resonance curve is of the form shown in Fig. 279. Since this curve is symmetrical about ω_m, the interval $\Delta\omega_m$ between values of ω_2 on either side of resonance giving the same value of $\mathcal{J}$ is twice the interval $\Delta\omega$, and we may write

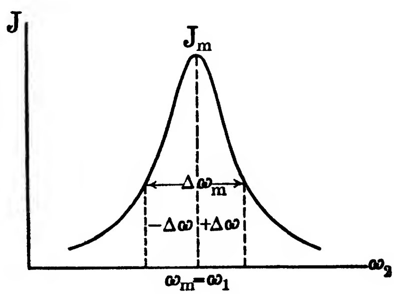


FIG. 279

$$\frac{\mathcal{J}}{\mathcal{J}_m} = \frac{(\alpha_1 + \alpha_2)^2}{(\alpha_1 + \alpha_2)^2 + \left(\frac{\Delta\omega_m}{2}\right)^2}. \tag{127–2}$$

Solving (127–2) for $\alpha_1 + \alpha_2$ we have

$$\alpha_1 + \alpha_2 = \frac{\Delta\omega_m}{2}\sqrt{\frac{\mathcal{J}}{\mathcal{J}_m - \mathcal{J}}}, \tag{127–3}$$

and, dividing by $\omega_m/2\pi$,

$$\delta_1 + \delta_2 = \pi \frac{\Delta\omega_m}{\omega_m} \sqrt{\frac{\mathcal{J}}{\mathcal{J}_m - \mathcal{J}}}. \tag{127-4}$$

Hence the procedure in making measurements consists in tuning the measuring circuit to resonance to obtain $\mathcal{J}_m$ and then detuning it to the same value $\mathcal{J}$ on either side. If the variable condenser used in the tuning is calibrated in terms of actual frequency ν we obtain $\Delta\omega_m/\omega_m = \Delta\nu_m/\nu_m$ directly. On the other hand, if the inductance is not known and the condenser is calibrated in terms of capacity we have

$$\frac{\Delta\omega_m}{\omega_m} = \frac{\dfrac{1}{\sqrt{L_2\left(C_m - \dfrac{\Delta C_m}{2}\right)}} - \dfrac{1}{\sqrt{L_2\left(C_m + \dfrac{\Delta C_m}{2}\right)}}}{\dfrac{1}{\sqrt{L_2 C_m}}} = \frac{1}{2}\frac{\Delta C_m}{C_m},$$

where C_m is the resonance setting of the condenser and ΔC_m is the capacity difference corresponding to $\Delta\omega_m$. It is convenient to choose $\mathcal{J} = \mathcal{J}_m/2$ since then (127–4) reduces to

$$\delta_1 + \delta_2 = \pi \frac{\Delta\omega_m}{\omega_m}. \tag{127-5}$$

Care must be taken to note whether the ammeter scale gives mean-square or root-mean-square currents. In the latter case the scale readings must be squared to obtain the $\mathcal{J}$'s.

Our measurements determine the sum of the logarithmic decrements of the source of oscillations and the auxiliary circuit, so one must be known in order to obtain the other. Usually δ_2 is known, the measuring circuit being in fact a wave-meter with a double scale, to give both frequency and logarithmic decrement. Such an instrument, often called a *decremeter*, provides a simple means of determining the logarithmic decrement of any source of damped oscillations.

Measurements with Undamped Waves. — In this case an adjustable circuit with a current indicator is loosely coupled to a

circuit containing a sinusoidal e.m.f. of angular frequency ω. Then from (90–8) and (126–16) the mean-square current in the adjustable circuit is of the form

$$\mathcal{J} = \tfrac{1}{2}i_{20}^2 = \frac{\tfrac{1}{2}M^2\omega^2\mathcal{E}_0^2}{(R_1^2 + X_1^2)(R_2^2 + X_2^2)},$$

and

$$\frac{\mathcal{J}}{\mathcal{J}_m} = \frac{R_2^2}{R_2^2 + X_2^2} = \frac{1}{1 + \left(\frac{L_2\omega}{R_2}\right)^2\left(1 - \frac{\omega_2^2}{\omega^2}\right)^2}. \qquad (127\text{–}6)$$

Now $L_2\omega/R_2 = \pi/\delta_2$ for the resonance adjustment where $\omega_2 = \omega \equiv \omega_m$. Also,

$$\left(1 - \frac{\omega_2^2}{\omega^2}\right)^2 = \left(\frac{\omega_m^2 - \omega_2^2}{\omega_m^2}\right)^2 = \left(\frac{[\omega_m + \omega_2][\omega_m - \omega_2]}{\omega_m^2}\right)^2$$
$$\doteq \left(\frac{2[\omega_m - \omega_2]}{\omega_m}\right)^2.$$

The resonance curve is the same as in the case of damped waves (Fig. 279) and $2[\omega_m - \omega_2] \equiv \Delta\omega_m$ is the frequency interval between equal values of $\mathcal{J}$ as before. Thus (127–6) becomes

$$\frac{\mathcal{J}}{\mathcal{J}_m} = \frac{1}{1 + \left(\frac{\pi}{\delta_2}\right)^2\left(\frac{\Delta\omega_m}{\omega_m}\right)^2},$$

and

$$\delta_2 = \pi\frac{\Delta\omega_m}{\omega_m}\sqrt{\frac{\mathcal{J}}{\mathcal{J}_m - \mathcal{J}}}. \qquad (127\text{–}7)$$

As (127–7) is like (127–4) the experimental procedure is the same with undamped waves as with damped waves. Here, however, we determine the logarithmic decrement of the adjustable circuit alone, so that this type of measurement is suited to the calibration of a decremeter or any similar device. It must be remembered that decrement determinations of this sort include the effect of the ammeter, which may be appreciable in low resistance circuits.

128. Filters. — We have investigated various simple a.c. networks in which a limited number of impedance elements are used. Let us now turn our attention to a general type of network in which numbers of similar impedance elements are assembled to form a recurrent structure. Networks of this sort are called *filters*, as they pass certain frequencies freely and stop others.

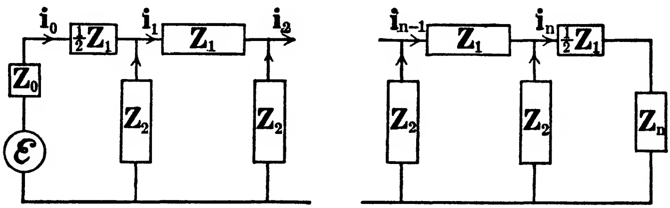


FIG. 280

Various forms of structure may be employed. We shall confine our discussion to the *ladder* type, illustrated in Fig. 280, which is one of the most common and most useful. Here there are a number of identical series elements of complex impedance $\mathbf{Z}_1$ and a number of shunt elements of complex impedance $\mathbf{Z}_2$. The input and output impedances are $\mathbf{Z}_0$ and $\mathbf{Z}_n$, respectively, and there is a complex applied e.m.f. $\boldsymbol{\mathcal{E}} = \mathcal{E}_0 e^{i\omega t}$, corresponding to an actual e.m.f. $\mathcal{E} = \mathcal{E}_0 \sin \omega t$. As shown, the filter proper ends with half series elements and is said to have *mid-series* terminations. We may think of the filter as composed of n T-sections (Fig. 281). Sometimes, however, it is more convenient to arrange the filter as illustrated

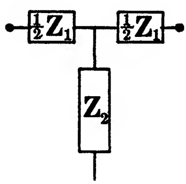


FIG. 281

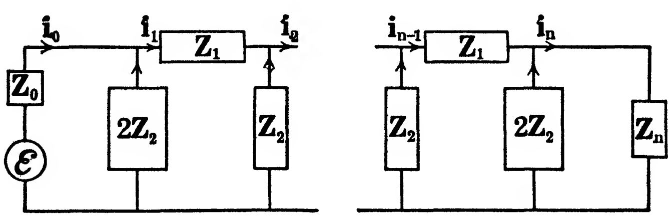


FIG. 282

in Fig. 282, with *mid-shunt* terminations. In this case we regard the filter as made up of $(n - 1)$ Π-sections (Fig. 283). The type of termination affects the values of the currents in the different sections of the filter for given input and output impedances, of course, but it does not affect the general frequency characteristics of the filter. Thus we need analyze only one case and we shall choose the mid-series.

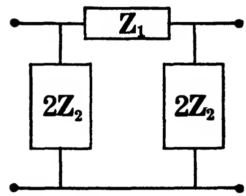


FIG. 283

Using Kirchhoff's laws in the complex form (art. 114), with positive directions of currents as indicated in Fig. 280, we have

$$\left.\begin{aligned} \mathbf{Z}_0\mathbf{i}_0 + \tfrac{1}{2}\mathbf{Z}_1\mathbf{i}_0 - \mathbf{Z}_2(\mathbf{i}_1 - \mathbf{i}_0) &= \mathcal{E}, \\ \mathbf{Z}_2(\mathbf{i}_1 - \mathbf{i}_0) + \mathbf{Z}_1\mathbf{i}_1 - \mathbf{Z}_2(\mathbf{i}_2 - \mathbf{i}_1) &= 0, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad & \cdot \\ \mathbf{Z}_2(\mathbf{i}_{n-1} - \mathbf{i}_{n-2}) + \mathbf{Z}_1\mathbf{i}_{n-1} - \mathbf{Z}_2(\mathbf{i}_n - \mathbf{i}_{n-1}) &= 0, \\ \mathbf{Z}_2(\mathbf{i}_n - \mathbf{i}_{n-1}) + \tfrac{1}{2}\mathbf{Z}_1\mathbf{i}_n + \mathbf{Z}_n\mathbf{i}_n &= 0. \end{aligned}\right\} \quad (128\text{–}1)$$

As we assume that a steady state exists, the currents have the form $\mathbf{i}_0 = \mathbf{i}_{00}e^{i\omega t}$, $\mathbf{i}_1 = \mathbf{i}_{10}e^{i\omega t}$ and so on. Hence, removing the factors $e^{i\omega t}$ and arranging the equations symmetrically, (128–1) becomes

$$\left.\begin{aligned} -\mathbf{Z}_2\mathbf{i}_{-10} + (\mathbf{Z}_1 + 2\mathbf{Z}_2)\mathbf{i}_{00} - \mathbf{Z}_2\mathbf{i}_{10} &= 0, \\ -\mathbf{Z}_2\mathbf{i}_{00} + (\mathbf{Z}_1 + 2\mathbf{Z}_2)\mathbf{i}_{10} - \mathbf{Z}_2\mathbf{i}_{20} &= 0, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad & \cdot \\ -\mathbf{Z}_2\mathbf{i}_{(n-2)0} + (\mathbf{Z}_1 + 2\mathbf{Z}_2)\mathbf{i}_{(n-1)0} - \mathbf{Z}_2\mathbf{i}_{n0} &= 0, \\ -\mathbf{Z}_2\mathbf{i}_{(n-1)0} + (\mathbf{Z}_1 + 2\mathbf{Z}_2)\mathbf{i}_{n0} - \mathbf{Z}_2\mathbf{i}_{(n+1)0} &= 0, \end{aligned}\right\} \quad (128\text{–}2)$$

where the quantities $\mathbf{i}_{-10}$ and $\mathbf{i}_{(n+1)0}$ are defined by

$$\left.\begin{aligned} \mathbf{i}_{-10} &= \frac{-(\mathbf{Z}_0 - \frac{1}{2}\mathbf{Z}_1 - \mathbf{Z}_2)\mathbf{i}_{00} + \mathcal{E}_0}{\mathbf{Z}_2}, \\ \mathbf{i}_{(n+1)0} &= \frac{-(\mathbf{Z}_n - \frac{1}{2}\mathbf{Z}_1 - \mathbf{Z}_2)\mathbf{i}_{n0}}{\mathbf{Z}_2}. \end{aligned}\right\} \quad (128\text{–}3)$$

The equations of (128–2) are of the form

$$-\mathbf{Z}_2\mathbf{i}_{(m-1)0} + (\mathbf{Z}_1 + 2\mathbf{Z}_2)\mathbf{i}_{m0} - \mathbf{Z}_2\mathbf{i}_{(m+1)0} = 0, \quad (128\text{–}4)$$

in which m takes all integral values from 0 to n, inclusive. We may expect a progressive change in the current as we pass from one section of the filter to the next, due to the recurrent structure. Let us, therefore, assume as a solution of (128–4)

$$\mathbf{i}_{m0} = \mathbf{A}e^{\Gamma m} + \mathbf{B}e^{-\Gamma m}, \tag{128–5}$$

where **A**, **B** and **Γ** are complex constants to be determined. Substituting (128–5) in (128–4) gives

$$= 0,$$

which reduces to

$$[(\mathbf{Z}_1 + 2\mathbf{Z}_2) - \mathbf{Z}_2(e^{\Gamma} + e^{-\Gamma})][\mathbf{A}e^{\Gamma m} + \mathbf{B}e^{-\Gamma m}] = 0.$$

Hence (128–5) is a solution of (128–4), provided only

$$e^{\Gamma} + e^{-\Gamma} = \frac{\mathbf{Z}_1 + 2\mathbf{Z}_2}{\mathbf{Z}_2}.$$

Here we may use hyperbolic functions to advantage. The hyperbolic sine (sinh) and the hyperbolic cosine (cosh) are defined by

$$\sinh x = \frac{e^x - e^-}{} \qquad \cosh x =$$

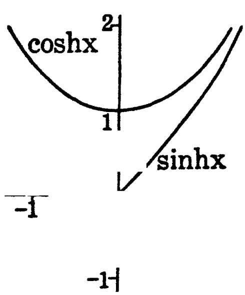


Fig. 284

Evidently $\sinh x$ is an odd function of x and $\cosh x$ is an even function. The graphs of these functions, assuming x real, are shown in Fig. 284. Note that $\cosh x$ is never less than unity, which is its value for $x = 0$. When the argument of the functions is imaginary we have

$$\sinh (iy) = \qquad i \sin y,$$

$$\cosh (iy) = \qquad = \cos y,$$

where y is real. It follows from these relations and from the general definition of the hyperbolic functions that

$$\begin{aligned}\cosh(x + iy) &= \cosh x \cosh(iy) + \sinh x \sinh(iy)\\ &= \cosh x \cos y + i \sinh x \sin y,\end{aligned}$$

a result which we shall find useful.

In terms of hyperbolic functions the equation determining $\mathbf{\Gamma}$ is then

$$\cosh \mathbf{\Gamma} = \frac{\mathbf{Z}_1 + 2\mathbf{Z}_2}{2\mathbf{Z}_2}. \tag{128–6}$$

The quantities **A** and **B** are evidently arbitrary constants depending on the boundary conditions, that is, on the values of $\mathbf{Z}_0$, $\mathbf{Z}_n$ and $\mathcal{E}_0$.

The frequency characteristics of the filter are controlled by $-\mathbf{\Gamma}$, which is called the *propagation constant*. For let us write

$$-\mathbf{\Gamma} \equiv \alpha + i\phi. \tag{128–7}$$

Then, since the actual currents are obtained by multiplying (128–5) by $e^{i\omega t}$ and taking the imaginary part, there is an attenuation factor $e^{-\alpha}$ introduced into the amplitude of the first term by each section as we move away from the input end, and similarly for the second term as we move toward the input end. Hence if α differs from zero for any frequency and the filter has an appreciable number of sections the current transmitted by the filter is effectively zero. On the other hand a current whose frequency is such that $\alpha = 0$ is freely transmitted. The quantity α is called the *attenuation constant*. Similarly ϕ is called the *phase constant*, since it gives the change of phase per section (measured as a lag) in the currents as we move along the filter.

The physical significance of (128–5) is now clear. The first term represents *space waves* of current traveling away from the input end, attenuated from section to section much as the time waves of current due to free oscillations in a single circuit are attenuated from instant to instant. The second term represents similar waves traveling back toward the input end arising from reflections at the output end.

There are two separate points of interest to be considered, the dependence of the frequency characteristics upon the filter's construction, that is, upon the nature of $\mathbf{Z}_1$ and $\mathbf{Z}_2$, and the dependence of the general transmission characteristics upon the terminating impedances $\mathbf{Z}_0$ and $\mathbf{Z}_n$. The first matter involves only relations (128–6) and (128–7) already obtained; the second involves the determination of **A** and **B**.

Frequency Characteristics. — We shall neglect resistances in the filter as they are always made small compared to associated reactances. Then $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are pure imaginaries and, in consequence, $\cosh \mathbf{\Gamma}$ is real. Expressing $\mathbf{\Gamma}$ in terms of α and ϕ,

$$\cosh \mathbf{\Gamma} = \cosh(-\mathbf{\Gamma}) = \cosh \alpha \cos \phi + i \sinh \alpha \sin \phi.$$

Here, as $\cosh \mathbf{\Gamma}$ is real, either α must be zero or ϕ must be an integral multiple of π. Hence, since $\cosh \alpha$ is never less than unity and $\cos \phi$ is never greater, we can distinguish three cases:

$$\left.\begin{array}{lll} -1 \leqq \cosh \mathbf{\Gamma} \leqq 1, & \alpha = 0, & \mathbf{\Gamma} = -i\phi; \\ \cosh \mathbf{\Gamma} > 1, & \phi = 0, & \mathbf{\Gamma} = -\alpha; \\ \cosh \mathbf{\Gamma} < -1, & \phi = \pm\pi, & \mathbf{\Gamma} = -\alpha \mp i\pi. \end{array}\right\} \quad (128\text{–}8)$$

In the frequency range corresponding to the first case currents are transmitted freely without attenuation, so we have a *pass band.* Frequency ranges corresponding to the other two cases are *stop bands.* Thus, in terms of $\mathbf{Z}_1$ and $\mathbf{Z}_2$, pass bands are given by

$$-1 \leqq \frac{\mathbf{Z}_1 + 2\mathbf{Z}_2}{2\mathbf{Z}_2} \leqq 1,$$

or

$$0 \leqq \left(-\frac{\mathbf{Z}_1}{\mathbf{Z}_2}\right) \leqq 4, \qquad (128\text{–}9)$$

and stop bands by all other ranges of values.

The simplest filters of the type we are considering are illustrated in Fig. 285. For the arrangement shown in (*a*),

$$\mathbf{Z}_1 = iL_1\omega, \qquad \mathbf{Z}_2 = -i\frac{1}{C_2\omega},$$

so that the pass band is given by

$$0 \leqq L_1C_2\omega^2 \leqq 4,$$

that is,

$$0 \leqq \omega \leqq \omega_c, \qquad \omega_c \equiv \frac{2}{\sqrt{L_1C_2}}. \tag{128–10}$$

As all frequencies from 0 to a critical or *cut-off* frequency given by $\omega_c = 2/\sqrt{L_1C_2}$ are passed without attenuation we have a *low-pass* filter.

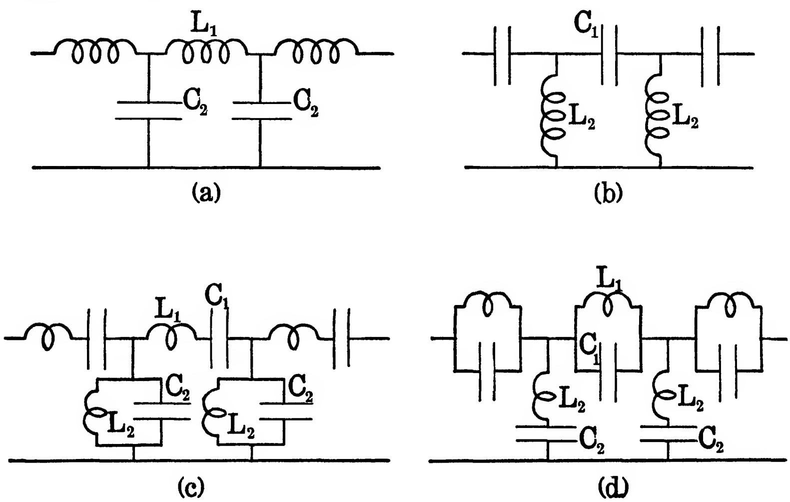


FIG. 285

In (b) the filter elements are

$$\mathbf{Z}_1 = -\,\mathrm{i}\frac{1}{C_1\omega}, \qquad \mathbf{Z}_2 = \mathrm{i}L_2\omega,$$

and the pass band is determined by

$$0 \leqq \frac{1}{L_2C_1\omega^2} \leqq 4,$$

that is,

$$\omega_c \leqq \omega \leqq \infty, \qquad \omega_c \equiv \frac{1}{2\sqrt{L_2C_1}}. \tag{128–11}$$

Here all frequencies above a critical frequency are passed without attenuation, so that we have a *high-pass* filter.

Arrangement (c) is slightly more complicated. As

$$\mathbf{Z}_1 = i\left(L_1\omega - \frac{1}{C_1\omega}\right), \qquad \mathbf{Z}_2 = i\frac{L_2\omega}{1 - L_2C_2\omega^2},$$

we have the inequality

$$0 \leqq \frac{(1 - L_1C_1\omega^2)(1 - L_2C_2\omega^2)}{L_2C_1\omega^2} \leqq 4,$$

which is satisfied by two ranges of values of ω. We usually prefer to bring the two ranges together, however, by taking $L_1C_1 = L_2C_2$. Then, setting

$$\omega_c \equiv \frac{1}{\sqrt{L_1C_1}} \equiv \frac{1}{\sqrt{L_2C_2}},$$

there is a single pass band given by

$$\left.\begin{aligned} &\omega_c' \leqq \omega \leqq \omega_c'', \\ &\omega_c' \equiv \omega_c\left[\sqrt{\sqrt{\frac{L_2C_1}{L_1C_2}} + 1} - \sqrt{\sqrt{\frac{L_2C_1}{L_1C_2}}}\right], \\ &\omega_c'' \equiv \omega_c\left[\sqrt{\sqrt{\frac{L_2C_1}{L_1C_2}} + 1} + \sqrt{\sqrt{\frac{L_2C_1}{L_1C_2}}}\right]. \end{aligned}\right\} \quad (128\text{–}12)$$

This filter is called a *band-pass* filter.

In (d) we have

$$\mathbf{Z}_1 = i\frac{L_1\omega}{1 - L_1C_1\omega^2}, \qquad \mathbf{Z}_2 = i\left(L_2\omega - \frac{1}{C_2\omega}\right).$$

The inequality

$$0 \leqq \frac{L_1C_2\omega^2}{(1 - L_1C_1\omega^2)(1 - L_2C_2\omega^2)} \leqq 4$$

is satisfied by two ranges of values of ω as in the previous case. Here we cannot bring the ranges together, but we will take $\omega_c \equiv 1/\sqrt{L_1C_1} \equiv 1/\sqrt{L_2C_2}$ as before, for simplicity. Then there are two pass bands:

$$0 \leqq \omega \leqq \omega_c', \qquad \omega_c' \equiv \frac{\omega_c}{4}\left[\sqrt{\sqrt{\frac{L_1C_2}{L_2C_1}} + 16} - \sqrt{\sqrt{\frac{L_1C_2}{L_2C_1}}}\right];$$

$$(128\text{–}13)$$

$$\omega_c'' \leqq \omega \leqq \infty, \qquad \omega_c'' \equiv \frac{\omega_c}{4}\left[\sqrt{\sqrt{\frac{L_1C_2}{L_2C_1}} + 16} + \sqrt{\sqrt{\frac{L_1C_2}{L_2C_1}}}\right].$$

As there is a single stop band corresponding to $\omega_c' < \omega < \omega_c''$ this filter is usually called a *band-stop* filter.

In addition to the location of the cut-off frequencies, it is often important to know how α varies with frequency in a stop band or how ϕ varies in a pass band. Using (128–6) to (128–8) inclusive, we see that in a stop band

$$\cosh \alpha = \pm \frac{}{2Z_2} \qquad (128\text{–}14)$$

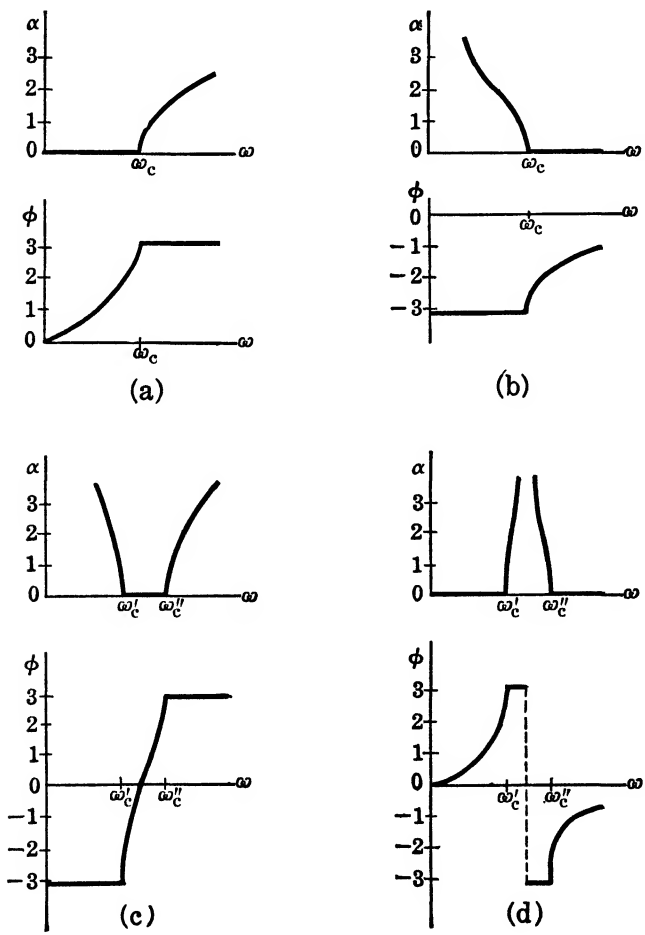


FIG. 286

the proper sign being used to make the right-hand side of the equation positive. Similarly, in a pass band,

$$\cos\phi = \frac{\mathbf{Z}_1 + 2\mathbf{Z}_2}{2\mathbf{Z}_2}. \qquad (128\text{–}15)$$

In Fig. 286 α and ϕ are shown as functions of the angular frequency for the four filters we have just discussed. To facilitate comparison the same value of ω_c is used in each case. The sign of ϕ is determined by including the resistance terms of $\mathbf{Z}_1$ and $\mathbf{Z}_2$ in the expanded form of (128–6) and observing that for $X_1 > 0$, $X_2 < 0$, ϕ is positive while for $X_1 < 0$, $X_2 > 0$, ϕ is negative. In other words ϕ always increases with frequency in pass bands.

Transmission Characteristics. — The values of **A** and **B** are most simply expressed in terms of the *characteristic* or *surge impedance* $\mathbf{Z}_K$ of the given filter. This is, by definition, the impedance of a filter of the type in question having an infinite number of sections. Evidently in this case waves due to reflection at the output end, represented by the second term in (128–5), vanish, the current in any section being independent of $\mathbf{Z}_n$, so that $\mathbf{Z}_K$ depends on the filter structure only. Referring to (128–5) we see that the complex current amplitudes are given by

$$\mathbf{i}_{m0} = \mathbf{i}_{00}e^{\Gamma m},$$

so that, from the first equation of (128–1),

$$[-\mathbf{Z}_2(e^{\Gamma} - 1) + \tfrac{1}{2}\mathbf{Z}_1 + \mathbf{Z}_0]\mathbf{i}_{00} = \mathcal{E}_0.$$

But

$$[\mathbf{Z}_K + \mathbf{Z}_0]\mathbf{i}_{00} = \mathcal{E}_0,$$

and hence

$$\mathbf{Z}_K = -\mathbf{Z}_2(e^{\Gamma} - 1) + \tfrac{1}{2}\mathbf{Z}_1 = -\mathbf{Z}_2\left(\frac{e^{\Gamma} - e^{-\Gamma}}{2}\right). \qquad (128\text{–}16)$$

Squaring this equation and using (128–6),

$$\left(\frac{\mathbf{Z}_K}{\mathbf{Z}_2}\right)^2 = \left(\frac{\mathbf{Z}_1 + 2\mathbf{Z}_2}{2\mathbf{Z}_2}\right)^2 - 1 = \frac{\mathbf{Z}_1}{\mathbf{Z}_2} + \frac{\mathbf{Z}_1^2}{4\mathbf{Z}_2^2},$$

or, finally,

$$\mathbf{Z}_K = \sqrt{\mathbf{Z}_1\mathbf{Z}_2 + \tfrac{1}{4}\mathbf{Z}_1^2}. \qquad (128\text{–}17)$$

We see that when we neglect the resistance terms in $\mathbf{Z}_1$ and $\mathbf{Z}_2$ as in (128–8), $\mathbf{Z}_K \equiv R_K$, a pure resistance, in the pass bands and $\mathbf{Z}_K \equiv iX_K$ in the stop bands.

Now, to determine $\mathbf{A}$ and $\mathbf{B}$ for the actual filter with a finite number of sections, substitute (128–5) in the first and the last equations of (128–1), obtaining

$$[-\mathbf{Z}_2(e^{\Gamma} - 1) + \tfrac{1}{2}\mathbf{Z}_1 + \mathbf{Z}_0]\mathbf{A} \\ + [-\mathbf{Z}_2(e^{-\Gamma} - 1) + \tfrac{1}{2}\mathbf{Z}_1 + \mathbf{Z}_0]\mathbf{B} = \mathcal{E}_0,$$

$$e^{\Gamma n}[-\mathbf{Z}_2(e^{-\Gamma} - 1) + \tfrac{1}{2}\mathbf{Z}_1 + \mathbf{Z}_n]\mathbf{A} \\ + e^{-\Gamma n}[-\mathbf{Z}_2(e^{\Gamma} - 1) + \tfrac{1}{2}\mathbf{Z}_1 + \mathbf{Z}_0]\mathbf{B} = 0.$$

From (128–16)

$$-\mathbf{Z}_2(e^{\Gamma} - 1) + \tfrac{1}{2}\mathbf{Z}_1 = \mathbf{Z}_K,$$

and, similarly,

$$-\mathbf{Z}_2(e^{-\Gamma} - 1) + \tfrac{1}{2}\mathbf{Z}_1 = -\mathbf{Z}_K.$$

Hence we have

$$\left.\begin{aligned} (\mathbf{Z}_K + \mathbf{Z}_0)\mathbf{A} - (\mathbf{Z}_K - \mathbf{Z}_0)\mathbf{B} &= \mathcal{E}_0, \\ -e^{\Gamma n}(\mathbf{Z}_K - \mathbf{Z}_n)\mathbf{A} + e^{-\Gamma n}(\mathbf{Z}_K + \mathbf{Z}_n)\mathbf{B} &= 0. \end{aligned}\right\} \qquad (128\text{–}18)$$

Now we will confine ourselves to the case where the output impedance *matches* the line, that is, $\mathbf{Z}_n = \mathbf{Z}_K$. Under this condition the filter with its terminating impedance behaves like an infinite filter. Reflections are absent, so that all the energy delivered to the filter at the input end is transmitted to $\mathbf{Z}_n$ and the filter operates at maximum efficiency. Setting $\mathbf{Z}_n = \mathbf{Z}_K$ in (128–18) gives

$$\mathbf{A} = \frac{\mathcal{E}_0}{\mathbf{Z}_K + \mathbf{Z}_0}, \qquad \mathbf{B} = 0,$$

and so

$$\mathbf{i}_{m0} = \frac{\mathcal{E}_0}{\mathbf{Z}_K + \mathbf{Z}_0}\, e^{\Gamma m}. \qquad (128\text{–}19)$$

In practice we are more interested in the input and output currents than in the intermediate currents. The completeness

of filtering is measured by

$$\frac{\mathbf{i}_{n0}}{\mathbf{i}_{00}} = e^{\Gamma n},$$

which is in general a function of the frequency. In a pass band the actual currents i_0 and i_n have the same amplitude but i_n lags behind i_0 in phase by an amount $n\phi$. In a stop band $i_{n0} = i_{00}e^{-n\alpha}$ and there is no phase difference other than a multiple of π. If the output impedance does not match the line the simple results just obtained do not hold. It must be observed that since $\mathbf{Z}_K$ is a function of the frequency no simple fixed output impedance can match the line over a range of frequencies.

Problem 128a. Show that the characteristic impedance $\mathbf{Z}_K'$ of a mid-shunt terminated filter (Fig. 282) is related to the corresponding mid-series characteristic impedance $\mathbf{Z}_K$ according to the equation $\mathbf{Z}_K'\mathbf{Z}_K = \mathbf{Z}_1\mathbf{Z}_2$.

Problem 128b. A low-pass filter with mid-series terminations is constructed of elements $L_1 = 1/\pi$ henry, $C_2 = 1/\pi$ microfarad. Find the cut-off frequency. Also calculate the characteristic impedance $\mathbf{Z}_K = R_K + \mathrm{i}X_K$ and plot as a function of frequency. Ans. $\nu_c = 1000$ cycle.

129. Lines. — The a.c. circuits and networks which we have considered up to this point are such that at any instant the current is the same all the way around the circuits or all the way along any branch of the networks. Let us now consider a case in which the current varies along the circuit at every instant as

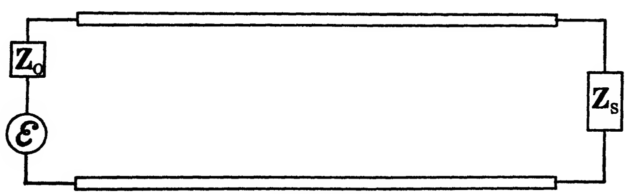


Fig. 287

well as varying in time at every point. Suppose we have two long parallel conductors of uniform, but not necessarily the same, cross-section (Fig. 287) connected to an impedance $\mathbf{Z}_0$ and e.m.f.

$\mathcal{E} = \mathcal{E}_0 e^{i\omega t}$ at one end and to an impedance $\mathbf{Z}_S$ at the other. The long conductors constitute a *line*, and $\mathbf{Z}_0$ and $\mathbf{Z}_S$ are its associated input and output impedances, respectively. Let the length of the line be S and the distance of any point on the line from the input end be s. If R and L are the resistance and inductance per unit length (including both sides of the line) and G and C are the leakage conductance and the capacity between conductors per unit length, we have for the series impedance per unit length $\mathbf{Z}_L = R + iL\omega$ and for the parallel or shunt impedance per unit length $\mathbf{Z}_C = 1/(G + iC\omega)$.

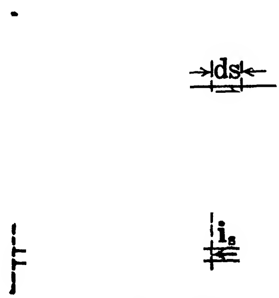


FIG. 288

Now consider an element ds of the line located at a distance s from the input end (Fig. 288). Denoting the complex current along the conductors and the complex e.m.f. between them at this point by $\mathbf{i}_s$ and $\mathcal{E}_s$ respectively we have

$$-\frac{\partial \mathbf{i}_s}{\partial s}\, ds = [(G\, ds) + i(C\, ds)\omega]\mathcal{E}_s = \frac{\mathcal{E}_s}{\mathbf{Z}_C}\, ds,$$

$$-\frac{\partial \mathcal{E}_s}{\partial s}\, ds = [(R\, ds) + i(L\, ds)\omega]\mathbf{i}_s = \mathbf{Z}_L \mathbf{i}_s ds.$$

Thus,

$$-\frac{\partial \mathbf{i}_s}{\partial s} = \frac{\mathcal{E}_s}{\mathbf{Z}_C}, \qquad -\frac{\partial \mathcal{E}_s}{\partial s} = \mathbf{Z}_L \mathbf{i}_s, \tag{129–1}$$

and, eliminating $\mathcal{E}_s$,

$$\frac{\partial^2 \mathbf{i}_s}{\partial s^2} - \frac{\mathbf{Z}_L}{\mathbf{Z}_C}\mathbf{i}_s = 0. \tag{129–2}$$

The frequency of the current is the same as that of the applied e.m.f., so that $\mathbf{i}_s = \mathbf{i}_{s0} e^{i\omega t}$. We may therefore reduce (129–1) and (129–2) to

$$-\frac{\partial \mathbf{i}_{s0}}{\partial s} = \frac{\mathcal{E}_{s0}}{\mathbf{Z}_C}, \qquad -\frac{\partial \mathcal{E}_{s0}}{\partial s} = \mathbf{Z}_L \mathbf{i}_{s0}, \tag{129–3}$$

and

$$\frac{\partial^2 \mathbf{i}_{s0}}{\partial s^2} - \frac{Z_L}{Z_C}\mathbf{i}_{s0} = 0, \tag{129-4}$$

respectively, and deal with the complex current amplitudes only.

The solution of (129–4) has the form

$$\mathbf{i}_{s0} = \mathbf{A}e^{\Gamma s} + \mathbf{B}e^{-\Gamma s}, \tag{129-5}$$

where $-\mathbf{\Gamma} \equiv \alpha + \mathrm{i}\phi$ is the *propagation constant* just as in the case of the filters discussed in article 128. Here α is the attenuation constant per unit length of line and ϕ is the corresponding phase constant per unit length. The first term of (129–5) represents waves of current traveling away from the input end of the line, continuously attenuated because of the factor $e^{-\alpha s}$. To see this express **A** in the exponential form $Ae^{-\mathrm{i}\alpha}$. Then on multiplying through by $e^{\mathrm{i}\omega t}$ the first term becomes $Ae^{-\alpha s}e^{\mathrm{i}(\omega t - \phi s - a)}$, the imaginary part of which, corresponding to an applied sine wave, is $Ae^{-\alpha s}\sin(\omega t - \phi s - a)$. This expression varies sinusoidally with the time at every point along the line and it also varies sinusoidally along the line at every instant of time. It therefore represents a wave traveling along the line, the direction of propagation being in the direction of increasing s, since increasing s requires increasing t to maintain a constant value of the total phase $(\omega t - \phi s - a)$. The *wave-length*, that is, the distance it is necessary to go along the line to find a phase difference of 2π radians at any instant, is given by

$$\lambda = \frac{2\pi}{\phi}, \tag{129-6}$$

while the speed v with which the waves travel is equal to the product of the wave-length by the frequency, that is,

$$= \frac{\omega}{\phi}. \tag{129-7}$$

The second term of (129–5) represents waves traveling back toward the input end, arising from reflections at the output end.

A wave of the sort under discussion is depicted in Fig. 289, where the real current i_s is shown as a function of distance for a particular instant of time. As time changes the wave crests

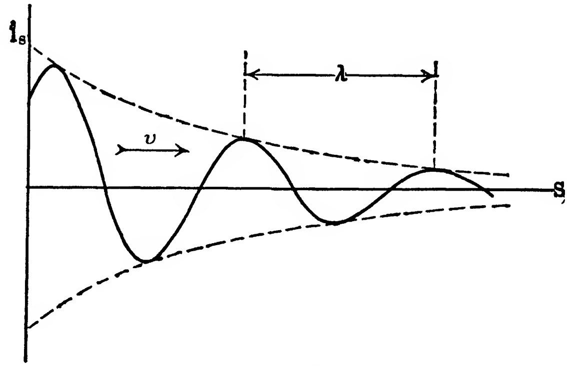


FIG. 289

move with the speed v keeping always between the broken line curves which represent the effect of the attenuation factor $e^{-\alpha s}$.

Substituting (129–5) in (129–4) shows us that

$$-\Gamma = \sqrt{\frac{Z_L}{Z_C}}. \tag{129–8}$$

Hence, squaring and substituting values,

$$(\alpha + i\phi)^2 = (R + iL\omega)(G + iC\omega),$$

or

$$\alpha^2 - \phi^2 = (RG - LC\omega^2),$$
$$2\alpha\phi = (GL + RC)\omega,$$

from which

$$\left.\begin{aligned} \alpha &= \sqrt{\frac{(RG - LC\omega^2) + \sqrt{(R^2 + L^2\omega^2)(G^2 + C^2\omega^2)}}{2}}, \\ \phi &= \sqrt{\frac{-(RG - LC\omega^2) + \sqrt{(R^2 + L^2\omega^2)(G^2 + C^2\omega^2)}}{2}}. \end{aligned}\right\} \tag{129–9}$$

Evidently both of these quantities have finite values over the entire range of frequencies between zero and infinity, so that

current waves of any frequency suffer a steady attenuation and lag of phase during propagation along the line. We shall discuss (129–9) in more detail below.

Next let us determine the arbitrary constants **A** and **B** in (129–5). These are best expressed in terms of the *characteristic impedance* $\mathbf{Z}_K$ of the line, that is, the impedance of an infinitely long line of the given type. For such a case (129–5) reduces to

$$\mathbf{i}_{s0} = \mathbf{i}_{00}e^{\Gamma s},$$

since the wave $\mathbf{B}e^{-\Gamma s}$ produced by reflection at the far end is absent. This may be substituted in the first equation of (129–3) giving

$$-\Gamma e^{\Gamma s}\mathbf{i}_{00} = \frac{\boldsymbol{\mathcal{E}}_{s0}}{\mathbf{Z}_C}.$$

Taking $s = 0$, we have, with the aid of (129–8),

$$\boldsymbol{\mathcal{E}}_{00} = \sqrt{\mathbf{Z}_L\mathbf{Z}_C}\,\mathbf{i}_{00}.$$

Hence, since $\mathbf{Z}_K = \boldsymbol{\mathcal{E}}_{00}/\mathbf{i}_{00}$,

$$\mathbf{Z}_K = \sqrt{\mathbf{Z}_L\mathbf{Z}_C}. \qquad (129\text{–}10)$$

Returning now to (129–5), substitution in the first equation of (129–3) gives

$$\mathbf{Z}_K(e^{\Gamma s}\mathbf{A} - e^{-\Gamma s}\mathbf{B}) = \boldsymbol{\mathcal{E}}_{s0}.$$

But

$$\boldsymbol{\mathcal{E}}_{00} = \mathcal{E}_0 - \mathbf{Z}_0\mathbf{i}_{00} = \mathcal{E}_0 - \mathbf{Z}_0(\mathbf{A} + \mathbf{B}),$$
$$\boldsymbol{\mathcal{E}}_{S0} = \mathbf{Z}_S\mathbf{i}_{S0} = \mathbf{Z}_S(e^{\Gamma S}\mathbf{A} + e^{-\Gamma S}\mathbf{B}),$$

so that

$$\left.\begin{aligned} (\mathbf{Z}_K + \mathbf{Z}_0)\mathbf{A} - (\mathbf{Z}_K - \mathbf{Z}_0)\mathbf{B} &= \mathcal{E}_0, \\ -e^{\Gamma S}(\mathbf{Z}_K - \mathbf{Z}_S)\mathbf{A} + e^{-\Gamma S}(\mathbf{Z}_K + \mathbf{Z}_S)\mathbf{B} &= 0. \end{aligned}\right\} \qquad (129\text{–}11)$$

Solving for **A** and **B** we find

$$\left.\begin{aligned} \mathbf{A} &= \frac{\mathcal{E}_0}{(\mathbf{Z}_K + \mathbf{Z}_0)}\,\frac{1}{(1 - e^{2\Gamma S}\mathbf{S}\mathbf{T})}, \\ \mathbf{B} &= \frac{\mathcal{E}_0}{(\mathbf{Z}_K + \mathbf{Z}_0)}\,\frac{e^{2\Gamma S}\mathbf{T}}{(1 - e^{2\Gamma S}\mathbf{S}\mathbf{T})}, \end{aligned}\right\} \qquad (129\text{–}12)$$

where

$$\mathbf{S} \equiv \frac{Z_K - Z_0}{Z_K + Z_0}, \qquad \mathbf{T} \equiv \frac{Z_K - Z_S}{Z_K + Z_S},$$

and, finally,

$$\mathbf{i}_{s0} = \frac{\mathcal{E}_0}{(Z_K + Z_0)} \frac{e^{\Gamma s} + e^{\Gamma(2S-s)}\mathbf{T}}{(1 - e^{2\Gamma S}\mathbf{ST})}. \qquad (129\text{–}13)$$

It is very instructive to expand (129–13) with the aid of the binomial theorem to the form

$$\mathbf{i}_{s0} = \frac{\mathcal{E}_0}{(Z_K + Z_0)} [e^{\Gamma s} + e^{\Gamma(2S-s)}\mathbf{T} + e^{\Gamma(2S+s)}\mathbf{ST} + e^{\Gamma(4S-s)}\mathbf{ST}^2 + e^{\Gamma(4S+s)}\mathbf{S}^2\mathbf{T}^2 + e^{\Gamma(6S-s)}\mathbf{S}^2\mathbf{T}^3 + \cdots]. \qquad (129\text{–}14)$$

Each term in this series has a definite physical significance. Since $\mathcal{E}_0/(Z_K + Z_0)$ gives the current flowing into an infinitely long line of the given type, the term $\frac{\mathcal{E}_0}{(Z_K + Z_0)} e^{\Gamma s}$ represents a current wave of characteristic initial amplitude which has traveled directly from the input end of the line to the point s. Similarly $\frac{\mathcal{E}_0}{(Z_K + Z_0)} e^{\Gamma(2S-s)}\mathbf{T}$ is the same current wave after it has traveled to the output end of the line, been reflected, and returned to the point s, a distance $2S - s$. $\mathbf{T}$, which evidently represents the effect of the reflection itself, is called the *output reflection coefficient.* The third term in the series gives the current wave after it has returned to the input end, been reflected with an *input reflection coefficient* $\mathbf{S}$ and traveled down the line to the point s again. Thus we may regard the current at any point along the line as due to a current wave $\mathcal{E}_0/(Z_K + Z_0)$ which enters at the input end and is then reflected back and forth across the point in question.

The point of view adopted above is helpful in considering the transmission characteristics of the line as a whole. For example, it calls to our attention the importance of matching impedances to obtain maximum efficiency of transmission. Thus when $Z_K = Z_0$, that is, $\mathbf{S} = 0$, maximum power is delivered to

the line, and when $\mathbf{Z}_S = \mathbf{Z}_K$, so that $\mathbf{T} = 0$, all power supplied to the line is in turn delivered to the load.

The actual current corresponding to an applied sine wave is obtained by multiplying (129–13) or (129–14) through by $e^{\iota\omega t}$ and taking the imaginary part, as usual. For example, consider the special case of a line matched at the input end but open at the output end, that is, $\mathbf{Z}_0 = \mathbf{Z}_K$ so that $\mathbf{S} = 0$, while $\mathbf{Z}_S = \infty$ so that $\mathbf{T} = -1$. For simplicity we will assume that R and G are negligible, in which case there is no attenuation and $\mathbf{Z}_K$ reduces to R_K, a pure resistance. Then (129–14) becomes

$$\mathbf{i}_{s0} = \frac{\mathcal{E}_0}{2R_K}\left[e^{-\iota\phi s} - e^{-\iota\phi(2S-s)}\right],$$

and the actual current is

$$i_s = \frac{\mathcal{E}_0}{2R_K}\left[\sin\omega\left(t - \frac{s}{v}\right) - \sin\omega\left(t - \frac{2S - s}{v}\right)\right],$$

ϕ having been replaced by ω/v in accord with (129–7). The first term represents the primary current wave and the second the reflected wave from the open end of the line. The combination of these waves results in *standing waves.* That is, there is a fixed amplitude of oscillation at every point on the line, the amplitude varying sinusoidally between zero and a maximum, $\mathcal{E}_0/R_K$, as we pass along the line. Since $\omega/v = 2\pi/\lambda$ the points of zero oscillation, called *nodes*, evidently occur at

$$s = S, \qquad S - \frac{\lambda}{2}, \qquad S - \lambda, \qquad \cdots,$$

while the points of maximum oscillation, called *loops*, occur at

$$s = S - \frac{\lambda}{4}, \qquad S - \frac{3\lambda}{4}, \qquad S - \frac{5\lambda}{4}, \qquad \cdots,$$

half-way between the nodes. Standing waves are usually shown graphically as in Fig. 290, the vertical distance between the curves representing the limiting magnitude of oscillation, that is, twice the amplitude, at each point. As the nodes and loops may

be located experimentally, standing waves afford a means of determining an unknown wave-length directly. The method is particularly valuable at very high frequencies where the wave-lengths are small.

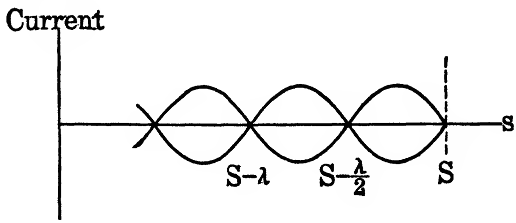


FIG. 290

With the line short-circuited at the output end instead of open, $\mathbf{T} = 1$, and the current is

$$i_s = \frac{\mathcal{E}_0}{2R_K}\left[\sin\omega\left(t - \frac{s}{v}\right) + \sin\omega\left(t - \frac{2S - s}{v}\right)\right].$$

In this case the standing waves are as shown in Fig. 291. Standing waves are, in fact, produced by any termination of the line other than $\mathbf{Z}_S = \mathbf{Z}_K$ and the distance between successive nodes or successive loops is always $\lambda/2$.

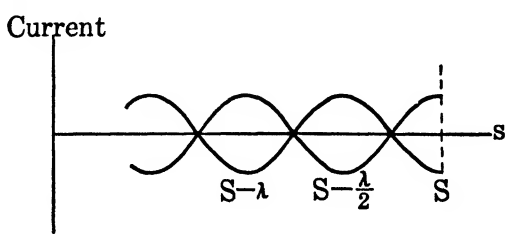


FIG. 291

Lines are used both for the transmission of power and for communication. In the former case only one frequency is involved and the primary requisite is small attenuation. In the latter case, on the other hand, it is usually necessary to transmit an entire band of frequencies and the essential requirement is that attenuation and speed of propagation should be

independent of frequency, in order to avoid distortion. Returning to (129–9) we see that, in general, this condition is not satisfied. If, however, we make

$$GL = RC, \tag{129–15}$$

then

$$\left.\begin{aligned} \alpha &= \sqrt{RG}, \\ v &= \frac{1}{\sqrt{LC}}, \end{aligned}\right\} \tag{129–16}$$

and the line is distortionless. Unfortunately it is not feasible, in practice, to satisfy (129–15) as it results in prohibitively high attenuation, so we must proceed otherwise.

Suppose that G is made negligible and that R is made small. Then

$$\alpha = \sqrt{\frac{LC\omega^2}{2}\left(\sqrt{1 + \frac{R^2}{L^2\omega^2}} - 1\right)}$$
$$= \sqrt{\frac{LC\omega^2}{2}\left(1 + \frac{1}{2}\frac{R^2}{L^2\omega^2} - \cdots - 1\right)},$$
$$\phi = \sqrt{\frac{LC\omega^2}{2}\left(\sqrt{1 + \frac{R^2}{L^2\omega^2}} + 1\right)}$$
$$= \sqrt{\frac{LC\omega^2}{2}\left(1 + \frac{1}{2}\frac{R^2}{L^2\omega^2} - \cdots + 1\right)}.$$

If, now, we can make $R^2/L^2\omega^2 \ll 1$, we have

$$\left.\begin{aligned} \alpha &= \frac{R}{2}\sqrt{\frac{C}{L}}, \\ v &= \frac{1}{\sqrt{LC}}, \end{aligned}\right\} \tag{129–17}$$

and the line is again distortionless. To make $R^2/L^2\omega^2$ sufficiently small it is necessary, in the audio frequency range, at least, to make L large, as R cannot economically be reduced beyond a certain point. There are two practicable ways in which large values of L may be obtained. The first consists of enclosing the line conductors in a thin sheath of some highly permeable sub-

stance such as permalloy or perminvar (art. 101). This method is very expensive to employ and its use is confined commercially to long distance submarine cables. The second way of increasing L depends on the insertion in the line at regular intervals of small inductance coils. While this procedure introduces a recurrent structure which causes the line to cut off at some critical frequency like a low-pass filter, nevertheless at lower frequencies it is approximately equivalent to increasing the inductance of the line uniformly. For good results several coils per wave-length must be used. This method is widely used for long distance telephone lines on land. A line whose inductance has been increased above the normal by either of the means just described is said to be *loaded.* Note that loading not only eliminates distortion but also reduces attenuation, a very desirable circumstance.

It remains, as a matter of some interest, to determine the magnitude of v for some actual lines. Confining ourselves to open wire lines, the capacity per unit length between two wires of circular cross-section of radius a, separated by a distance b large compared to a, is given in e.m.u. by

$$C = \frac{1}{4c^2 \log \frac{b}{a}},$$

from (32–3), using the conversion table (p. 442) to change from e.s.u. to e.m.u.

The corresponding inductance per unit length of line, assuming there is no loading, is found by calculating the flux which links a unit length of the line when a unit current is flowing. Taking the origin of rectangular coordinates half-way between the wires as indicated in Fig. 292 we

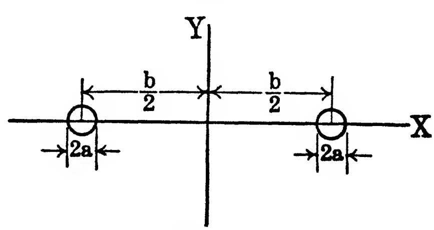


FIG. 292

have

$$L = \int_{-(b/2)+a}^{(b/2)-a} \left(\frac{2}{\frac{b}{2} - x} + \frac{2}{\frac{b}{2} + x} \right) dx$$

$$\left| 2 \log \frac{\frac{b}{2} + x}{\frac{b}{2} - x} \right|_{-(b/2)+a}^{(b/2)-a} = 4 \log \frac{b}{a},$$

since b is large compared with a. Thus, in the ideal case of a resistanceless line,

$$v = \frac{1}{\sqrt{LC}} = c, \tag{129-18}$$

regardless of the size of the wires or the spacing. As has been pointed out in Chapter XII, c is equal to the speed of light. It is in fact the speed with which any electromagnetic disturbance travels through free space. For an actual unloaded line where $R^2/L^2\omega^2$ is not negligible v is less than c and it depends both on the structure of line and on the frequency, but the discrepancy is not very great when large gauge low resistance wires are used. Loading the line to the extent necessary for satisfactory voice transmission reduces v to about one-third c or less. That is, the speed on an average open wire loaded telephone line is of the order of $(10)^5$ kilometers per second. The corresponding wave-length at, for example, a thousand cycles is a hundred kilometers.

Problem 129a. Show that the complex amplitude of the e.m.f. at any point on a line is

$$\mathcal{E}_{s0} = \mathcal{E}_0 \frac{\mathbf{Z}_K}{(\mathbf{Z}_K + \mathbf{Z}_0)} \frac{e^{\Gamma s} - e^{\Gamma(2S-s)}\,\mathbf{T}}{(1 - e^{2\Gamma S}\,\mathbf{ST})}.$$

What are the input and output reflection coefficients? Ans. $-\mathbf{S}$, $-\mathbf{T}$.

Problem 129b. Find the relation between e.m.f. and current in standing waves. Ans. Loops of e.m.f. occur at current nodes, and *vice versa.*

Problem 129c. Show that the characteristic impedance of a loaded line is practically a pure resistance given by $R_K = \sqrt{L/C}$.

Problem 129d. Show that if a line of characteristic impedance $\mathbf{Z}_K$ one-quarter of a wave-length long is terminated by an output impedance $\mathbf{Z}_S$, the total impedance measured at the input end of the line is $\mathbf{Z}_K^2/\mathbf{Z}_S$. Assume the effect of attenuation to be negligible.

CHAPTER XVI

ELECTROMAGNETIC WAVES

130. Electromagnetic Equations. — The fundamental equations describing the electromagnetic field produced by charges at rest or in motion relative to the observer may be taken as (1) Gauss' law (14–4) specifying the flux of displacement through a closed surface s surrounding free charge of density ρ, (2) Gauss' law (43–3) for the flux of induction in the corresponding magnetic case, (3) Ampère's law (85–6) for the magnetic field due to a current, and (4) Faraday's law (85–5) for the electric field produced by a changing magnetic flux. In addition we need relations such as (15–3) between D and E and (43–5) between B and H.

In the investigation of electromagnetic fields to be undertaken in this chapter we shall use all these relations; consequently we must write them all in the same system of units. We shall employ Heaviside-Lorentz units since they exhibit the symmetry of the electromagnetic equations and enable us to dispense with irrelevant numerical factors. This system of units is more generally used in works on electromagnetic theory than either of the older c.g.s. systems.

Gauss' law for electrostatics has already been expressed in h.l.u. in (108–3), Gauss' law for magnetostatics is

$$\int_s \mathbf{B}_l \cdot d\mathbf{s} = 0,$$

and Ampère's and Faraday's laws are given in h.l.u. in article 108. So, dropping the subscript l, we have altogether

$$\int_s \mathbf{D}\cdot d\mathbf{s} = \int_\tau \rho\, d\tau, \qquad (a) \qquad \int_s \mathbf{B}\cdot d\mathbf{s} = 0, \qquad (b)$$

$$\oint \mathbf{H}\cdot d\mathbf{l} = \frac{1}{c}\int_s \mathbf{j}\cdot d\mathbf{s}, \qquad (c) \qquad \oint \mathbf{E}\cdot d\mathbf{l} = -\frac{1}{c}\int_s \frac{\partial \mathbf{B}}{\partial t}\cdot d\mathbf{s}, \qquad (d)$$

in Heaviside-Lorentz units. In (*a*) and (*b*) the surface integral on the left is taken over a *closed* surface, whereas the surface integral on the right of (*c*) and (*d*) is taken over an *open* surface bounded by the closed curve along which the line integral on the left is evaluated.

Maxwell's great contribution to electromagnetic theory had its origin in his recognition of the fact that these laws fail to satisfy the equation of continuity. This equation states that the time rate of increase of charge inside a volume τ is equal to the excess of the rate at which charge flows into τ through the boundary surface over that at which it flows out. Physically it is equivalent to the statement that charge can neither be created nor destroyed. For if charge is indestructible, the only way in which it can accumulate in a region τ is by more charge entering the region than leaving it.

Maxwell traced the difficulty to Ampère's law, which he showed to be incomplete in the form (*c*). In fact, he found it necessary to add a term to the right-hand side of this equation involving the integral over s of the time rate of change of displacement. Without this added term the electromagnetic theory of light would have been impossible. We shall now examine the reasons which led Maxwell to modify Ampère's law, although we shall develop them from a somewhat different point of view than that adopted by him in his original investigation.

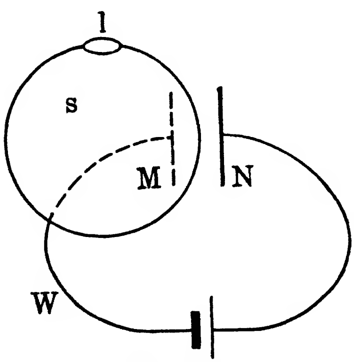


FIG. 293

If we apply Ampère's law (*c*) to a surface s (Fig. 293) which is closed except for a small opening, the line integral on the left is taken over only the boundary l of the opening. If we make the hole smaller and smaller, the periphery becomes shorter and shorter, until finally, when the surface is completely closed, the left-hand member of

(c) reduces to zero. So, for a *closed* surface s, Ampère's law becomes

$$\int_s \mathbf{j} \cdot d\mathbf{s} = 0. \tag{130–1}$$

In words, this relation states that the current entering through one part of s must be exactly compensated by the current leaving through other parts of the surface. But this stringent requirement is not at all in accord with experience. If, for instance, the surface s surrounds one plate M of a condenser, we can certainly charge up M by a current entering s through the wire W without any compensating conduction current flowing out through the surface. Consequently we conclude that Ampère's law cannot be properly expressed in the form (c).

To correct Ampère's law, we will calculate the time rate of increase of charge in a volume τ from (a). If γ is the angle between $\mathbf{D}$ and $d\mathbf{s}$,

$$\int_s \mathbf{D} \cdot d\mathbf{s} = \int_s D \cos \gamma ds,$$

and the rate of increase of charge is

$$\int_\tau \frac{\partial \rho}{\partial t} d\tau = \int_s \frac{\partial}{\partial t}(D \cos \gamma) ds.$$

But if α is the angle which the vector $\dfrac{\partial \mathbf{D}}{\partial t}$ makes with $d\mathbf{s}$,

$$\partial t \quad \cos \alpha = \frac{\partial}{\partial t}(D \cos \gamma),$$

as in (85–3), and

$$\int_s \frac{\partial}{\partial t}(D \cos \gamma) ds = \int_s \left| \frac{\partial \mathbf{D}}{\partial t} \right| \cos \alpha ds = \int_s \frac{\partial \mathbf{D}}{\partial t} \cdot d\mathbf{s}.$$

So

$$\int_\tau \frac{\partial \rho}{\partial t} d\tau = \int_s \frac{\partial \mathbf{D}}{\partial t} \cdot d\mathbf{s} \qquad ($$

represents the rate at which charge accumulates inside the closed surface s.

But the excess of the rate at which charge flows into τ over that at which it leaves is $-\int \mathbf{j}\cdot d\mathbf{s}$ integrated over the closed surface s bounding τ. So, if the equation of continuity is to hold,

$$\int_s \left(\mathbf{j} + \frac{\partial \mathbf{D}}{\partial t}\right) \cdot d\mathbf{s} = 0 \qquad (130\text{–}3)$$

must be valid instead of (130–1) for any closed surface s. Consequently we must replace the current density in Ampère's law by the integrand of (130–3), getting

$$\oint \mathbf{H}\cdot d\mathbf{l} = \frac{1}{c}\int_s \left(\mathbf{j} + \frac{\partial \mathbf{D}}{\partial t}\right) \cdot d\mathbf{s} \qquad (130\text{–}4)$$

as the complete expression for the magnetomotive force around the periphery of an open surface s.

Effectively we have added to the conduction or convection current $\mathbf{j}$ an additional current equal in density to the time rate of increase of displacement. Maxwell called this a *displacement current.* In part it represents a true motion of electrical charges. For the displacement in a dielectric is

$$\mathbf{D} = \mathbf{E} + \mathbf{P}$$

in h.l.u., and

$$\frac{\partial \mathbf{D}}{\partial t} = \frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \mathbf{P}}{\partial t}.$$

Now if ρ_1 is the charge per unit volume which suffers a displacement $\mathbf{R}$ when the medium is polarized (art. 13), $\mathbf{P} = \rho_1\mathbf{R}$, and

$$\frac{\partial \mathbf{P}}{\partial t} = \rho_1 \frac{\partial \mathbf{R}}{\partial t}$$

represents the current passing through a unit cross-section due to the velocity of the displaced charge. This is a true electrical current, due, like the current in a conductor, to a flow of electricity.

It should be noted that when we take account of displacement currents as well as of conduction currents all currents flow around *closed* circuits. For (130–3) tells us that any conduction current

leading into the region surrounded by a closed surface s must be compensated by an equal displacement current flowing out. Thus, in the case illustrated in Fig. 293, the current passing through the wire W to the plate M of the condenser is accompanied by a displacement current flowing through the dielectric from the plate M to the plate N. The conduction current entering through the surface s is exactly compensated by the displacement current passing out. The net result is that the *same* current—either in the form of a conduction current or of a displacement current—flows through every cross-section of the complete circuit $WMNW$.

In the case of a steady conduction current the added term in (130–4) vanishes, and in the case of slowly varying currents it is of importance only between the plates of a condenser in the circuit. So the calculations of the magnetic field produced by a current which were made in Chapter VII have not been invalidated by Maxwell's revision of Ampère's law.

In a region in which no conduction or convection currents are present, (130–4) becomes

$$\oint \mathbf{H}\cdot d\mathbf{l} = \frac{1}{c}\int_s \frac{\partial \mathbf{D}}{\partial t}\cdot d\mathbf{s}. \qquad (130\text{–}5)$$

This equation states that the magnetomotive force around any closed circuit is proportional to the time rate of change of flux of displacement through the circuit, just as Faraday's law states that the induced electromotive force around a circuit is proportional to the time rate of change of flux of induction through the circuit. Why, then, is this equation not as susceptible to direct experimental verification as Faraday's law? The reason is that there exists in nature no conductor of magnetism and therefore no sensitive method of detecting a magnetomotive force. Were it not that metals are excellent conductors of electricity Faraday would never have been able to detect so simply the electromotive force induced in a circuit by a varying flux of induction.

In addition to the four equations which have been under discussion we need a relation specifying the force exerted on a

charge by the electric and magnetic fields through which it is moving. This information is supplied by (76–2). Converting from the e.m.u. there used to h.l.u. we have the force per unit volume

$$\mathbf{F}_\tau = \rho\left(\mathbf{E} + \frac{1}{c}\mathbf{v} \times \mathbf{B}\right)$$

on a charge of volume density ρ moving with velocity $\mathbf{v}$ relative to the observer.

If we write $\rho\mathbf{v}$ in place of $\mathbf{j}$ for the current density and denote partial differentiation with respect to the time by a dot placed over the quantity involved, we have the complete set of electromagnetic equations:

$$\int_s \mathbf{D}\cdot d\mathbf{s} = \int_\tau \rho d\tau, \tag{1}$$

$$\int_s \mathbf{B}\cdot d\mathbf{s} = 0, \tag{2}$$

$$\oint \mathbf{H}\cdot d\mathbf{l} = \frac{1}{c}\int_s (\dot{\mathbf{D}} + \rho\mathbf{v})\cdot d\mathbf{s}, \tag{3}$$

$$\oint \mathbf{E}\cdot d\mathbf{l} = -\frac{1}{c}\int_s \dot{\mathbf{B}}\cdot d\mathbf{s}, \tag{4}$$

$$\mathbf{F}_\tau = \rho\left(\mathbf{E} + \frac{1}{c}\mathbf{v} \times \mathbf{B}\right), \tag{5}$$

$$\mathbf{D} = \mathbf{E} + \mathbf{P} = \kappa\mathbf{E}, \tag{6}$$

$$\mathbf{B} = \mathbf{H} + \mathbf{I} = \mu\mathbf{H}. \tag{7}$$

While the first five equations and the relations $\mathbf{D}=\mathbf{E}+\mathbf{P}$, $\mathbf{B} = \mathbf{H} + \mathbf{I}$ are quite general, the relation $\mathbf{D} = \kappa\mathbf{E}$ holds only for isotropic dielectrics and $\mathbf{B} = \mu\mathbf{H}$ only for isotropic paramagnetic or diamagnetic media. The first four equations, known as the *field equations*, together with (6) and (7), suffice to determine uniquely the electric and magnetic fields produced by an arbitrarily assigned distribution of charge density ρ and current density $\rho\mathbf{v}$. The *force equation* (5), on the other hand, specifies

the force per unit volume on charge of density ρ moving with velocity $\mathbf{v}$ through an electric field $\mathbf{E}$ and a magnetic field $\mathbf{B}$.

Problem 130a. A parallel plate condenser consists of two circular plates of radius a. It is connected to an alternating e.m.f. so that the charge in h.l.u. is $q = q_0 \sin \omega t$. Neglecting edge effect, find the magnetic intensity between the plates in terms of the distance r from the axis.

$$\text{Ans. } H = \frac{\omega q_0}{2\pi c a^2} r \cos \omega t.$$

Problem 130b. Give the answer to 130*a* in e.m.u.

$$\text{Ans. } H_m = \frac{2\omega (q_0)_m}{a^2} r \cos \omega t.$$

131. Differential Form of Field Equations. — Before we can deduce the wave equation and investigate the properties of electromagnetic waves we must express the field equations (1) to (4) of the previous article in differential form.

Equations (1) and (2) are easily transformed. Consider the small rectangular parallelopiped (Fig. 294) of dimensions dx, dy, dz. The flux of displacement through the two faces perpendicular to the X axis is

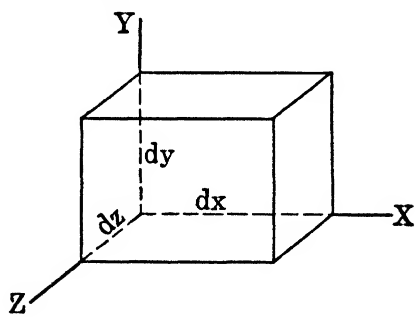


FIG. 294

$$-D_x dydz + \left(D_x + \frac{\partial D_x}{\partial x} dx \right) dydz = \frac{\partial D_x}{\partial x} dxdydz.$$

Adding similar expressions for the flux through the pairs of faces perpendicular to the Y and Z axes respectively, we have for the total outward flux

$$\left(\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} \right) dxdydz,$$

and equating this to the charge $\rho dxdydz$ inside the parallelopiped we get

$$\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} = \rho \tag{1'}$$

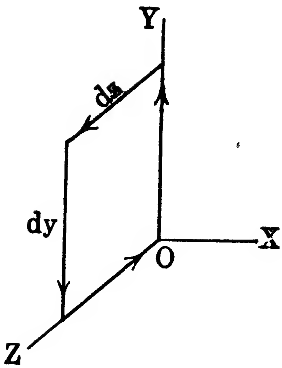


FIG. 295

for the differential form of (1). Similarly (2) becomes

$$\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0. \qquad (2')$$

To obtain Ampère's law in differential form we shall apply (3) to a small rectangular circuit (Fig. 295) of dimensions dy, dz lying in the YZ plane. Then, denoting the components of magnetic intensity at O by H_x, H_y, H_z,

$$H_y dy + \left(H_z + \frac{\partial H_z}{\partial y} dy \right) dz - \left(H_y + \frac{\partial H_y}{\partial z} dz \right) dy - H_z dz = \frac{1}{c} (\dot{D}_x + \rho v_x) dydz,$$

or

$$\left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} \right) dydz = \frac{1}{c} (\dot{D}_x + \rho v_x) dydz.$$

Dividing by the area $dydz$ of the circuit and writing similar expressions for the cases where the circuit lies in the ZX and XY planes,

$$\left.\begin{aligned} \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} &= \frac{1}{c} (\dot{D}_x + \rho v_x), \\ \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} &= \frac{1}{c} (\dot{D}_y + \rho v_y), \\ \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} &= \frac{1}{c} (\dot{D}_z + \rho v_z). \end{aligned}\right\} \quad (3')$$

From Faraday's law (4) we obtain in exactly similar fashion

$$\left.\begin{aligned} \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} &= -\frac{1}{c} \dot{B}_x, \\ \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} &= -\frac{1}{c} \dot{B}_y, \\ \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} &= -\frac{1}{c} \dot{B}_z. \end{aligned}\right\} \quad (4')$$

It may be noted that Ampère's law and Faraday's law each leads to *three* scalar equations, indicating that these laws are vector relations. The vector whose three components are given on the left-hand side of (3′) is known as the *curl* of **H**. Similarly on the left-hand side of (4′) we have the *curl* of **E**.

132. Electromagnetic Waves. — In a homogeneous isotropic medium containing no free charge the field equations (1′) to (4′) of article 131 become

$$\frac{\partial E_x}{\partial x}+\frac{\partial E_y}{\partial y}+\frac{\partial E_z}{\partial z}=0, \quad (1'') \qquad \frac{\partial H_x}{\partial x}+\frac{\partial H_y}{\partial y}+\frac{\partial H_z}{\partial z}=0, \quad (2'')$$

$$\left.\begin{aligned}\frac{\partial H_z}{\partial y}-\frac{\partial H_y}{\partial z}&=\frac{\kappa}{c}\dot{E}_x,\\ \frac{\partial H_x}{\partial z}-\frac{\partial H_z}{\partial x}&=\frac{\kappa}{c}\dot{E}_y,\\ \frac{\partial H_y}{\partial x}-\frac{\partial H_x}{\partial y}&=\frac{\kappa}{c}\dot{E}_z,\end{aligned}\right\}(3'') \qquad \left.\begin{aligned}\frac{\partial E_z}{\partial y}-\frac{\partial E_y}{\partial z}&=-\frac{\mu}{c}\dot{H}_x,\\ \frac{\partial E_x}{\partial z}-\frac{\partial E_z}{\partial x}&=-\frac{\mu}{c}\dot{H}_y,\\ \frac{\partial E_y}{\partial x}-\frac{\partial E_x}{\partial y}&=-\frac{\mu}{c}\dot{H}_z,\end{aligned}\right\}(4'')$$

where we have put κE for D and μH for B.

The first step in the deduction of the wave equation consists in eliminating the components of **H** from (3″) and (4″). To accomplish this, differentiate the first of the three equations (3″) with respect to the time, getting

$$\frac{\partial \dot{H}_z}{\partial y}-\frac{\partial \dot{H}_y}{\partial z}=\frac{\kappa}{c}\ddot{E}_x.$$

Next differentiate the second of (4″) with respect to z and the third with respect to y. This gives

$$\frac{\partial \dot{H}_y}{\partial z}=-\frac{c}{\mu}\left(\frac{\partial^2 E_x}{\partial z^2}-\frac{\partial^2 E_z}{\partial x\partial z}\right),$$

$$-\frac{\partial \dot{H}_z}{\partial y}=-\frac{c}{\mu}\left(\frac{\partial^2 E_x}{\partial y^2}-\frac{\partial^2 E_y}{\partial x\partial y}\right).$$

Adding the last three equations,

$$\frac{\partial^2 E_x}{\partial y^2}+\frac{\partial^2 E_x}{\partial z^2}-\frac{\partial}{\partial x}\left(\frac{\partial E_y}{\partial y}+\frac{\partial E_z}{\partial z}\right)=\frac{\kappa\mu}{c^2}\ddot{E}_x$$

Using (1″) to eliminate the terms in E_y and E_z and writing similar equations in the remaining components of **E** we have

$$\left.\begin{aligned}\frac{\partial^2 E_x}{\partial x^2}+\frac{\partial^2 E_x}{\partial y^2}+\frac{\partial^2 E_x}{\partial z^2}&=\frac{\kappa\mu}{c^2}\frac{\partial^2 E_x}{\partial t^2},\\ \frac{\partial^2 E_y}{\partial x^2}+\frac{\partial^2 E_y}{\partial y^2}+\frac{\partial^2 E_y}{\partial z^2}&=\frac{\kappa\mu}{c^2}\frac{\partial^2 E_y}{\partial t^2},\\ \frac{\partial^2 E_z}{\partial x^2}+\frac{\partial^2 E_z}{\partial y^2}+\frac{\partial^2 E_z}{\partial z^2}&=\frac{\kappa\mu}{c^2}\frac{\partial^2 E_z}{\partial t^2}.\end{aligned}\right\} \tag{132–1}$$

This set of relations is known as the *wave equation*. If we had eliminated **E** from the field equations instead of **H** we would have obtained an exactly similar set of equations in the components of **H**, namely,

$$\left.\begin{aligned}\frac{\partial^2 H_x}{\partial x^2}+\frac{\partial^2 H_x}{\partial y^2}+\frac{\partial^2 H_x}{\partial z^2}&=\frac{\kappa\mu}{c^2}\frac{\partial^2 H_x}{\partial t^2},\\ \frac{\partial^2 H_y}{\partial x^2}+\frac{\partial^2 H_y}{\partial y^2}+\frac{\partial^2 H_y}{\partial z^2}&=\frac{\kappa\mu}{c^2}\frac{\partial^2 H_y}{\partial t^2},\\ \frac{\partial^2 H_z}{\partial x^2}+\frac{\partial^2 H_z}{\partial y^2}+\frac{\partial^2 H_z}{\partial z^2}&=\frac{\kappa\mu}{c^2}\frac{\partial^2 H_z}{\partial t^2}.\end{aligned}\right\} \tag{132–2}$$

Let us consider the case where the components of **E** and **H** are functions of x and t only. Then (132–1) reduces to

$$\left.\begin{aligned}\frac{\partial^2 E_x}{\partial x^2}&=\frac{1}{v^2}\frac{\partial^2 E_x}{\partial t^2},\\ \frac{\partial^2 E_y}{\partial x^2}&=\frac{1}{v^2}\frac{\partial^2 E_y}{\partial t^2},\\ \frac{\partial^2 E_z}{\partial x^2}&=\frac{1}{v^2}\frac{\partial^2 E_z}{\partial t^2},\end{aligned}\right\} \tag{132–3}$$

where we have put

$$v\equiv\frac{c}{\sqrt{\kappa\mu}}. \tag{132–4}$$

A similar set of equations holds for the components of **H**.

Consider the first of the three equations (132–3). Its general solution is

$$E_x=f_1(x-vt)+f_2(x+vt),$$

where f_1 and f_2 are arbitrary functions, as can be verified at once by substitution in the differential equation. Now if we increase t by τ and x by $v\tau$,

$$+ v\tau] - v[t + \tau]) = f_1(x - vt),$$

showing that the function f_1 has the same value at the point $x + v\tau$ at the time $t + \tau$ as at x at the time t. Consequently this function represents a wave advancing in the positive X direction with the *phase velocity* v. Similarly the identity

$$f_2([x - v\tau] + v[t + \tau]) = f_2(x + vt)$$

shows that f_2 represents a wave traveling in the negative X direction with the same phase velocity. In both cases E_x has the same value at a given instant at all points in any plane perpendicular to the X axis. Therefore the wave is *plane*, the wave-front being infinite in extent.

If, now, we confine our attention to a plane wave moving in the positive X direction, the wave equation gives

$$\begin{aligned} E_x &= f(x - vt), & H_x &= i(x - vt), \\ E_y &= g(x - vt), & H_y &= j(x - vt), \\ E_z &= h(x - vt), & H_z &= k(x - vt). \end{aligned}$$

Substituting in (1″) and (2″) we find

$$\frac{\partial f}{\partial x} = 0, \qquad \frac{\partial i}{\partial x} = 0.$$

Therefore E_x and H_x are constants in both x ant t. As we are interested only in the variable part of the field, and not in electrostatic and magnetostatic fields which may be superposed on the electromagnetic wave, we shall take these constants to be zero. Consequently **E** and **H** are both at right angles to the direction of propagation. The wave, then, is transverse and in this respect has the characteristics of light.

The remaining components of **E** and **H** are not independent since they must satisfy the field equations (3″) and (4″). Denoting the derivative of a function with respect to its argument

by a prime, the second and third equations of (3″) give

$$-k' = -\frac{v\kappa}{c}g', \qquad j' = -\frac{v\kappa}{c}h',$$

or

$$k = \sqrt{\frac{\kappa}{\mu}}g, \qquad j = -\sqrt{\frac{\kappa}{\mu}}h,$$

except for a constant of integration which represents merely a superposed static field. The same relations might have been obtained from (4″).

Hence we have altogether

$$\left.\begin{aligned} E_x &= 0, & H_x &= 0, \\ E_y &= g(x - vt), & H_y &= -\sqrt{\frac{\kappa}{\mu}}h(x - vt), \\ E_z &= h(x - vt), & H_z &= \sqrt{\frac{\kappa}{\mu}}g(x - vt). \end{aligned}\right\} \quad (132\text{–}5)$$

If θ is the angle between **E** and **H**,

$$EH\cos\theta = \mathbf{E}\cdot\mathbf{H} = \sqrt{\frac{\kappa}{\mu}}\{-gh + gh\} = 0.$$

Consequently $\theta = \pi/2$, **E** and **H** lying in the wave front at right angles to each other.

We can consider that (132–5) represents two waves, each corresponding to one of the two possible states of polarization of a transverse wave. One of these waves, determined by the function $g(x - vt)$, consists of an electric field g in the Y direction and a magnetic field $\sqrt{\frac{\kappa}{\mu}}g$ in the Z direction; the other, specified by the function $h(x - vt)$, consists of an electric field h in the Z direction and a magnetic field $\sqrt{\frac{\kappa}{\mu}}h$ in the negative Y direction. In each wave the electric and magnetic fields are in phase and so oriented that a right-handed screw, rotated from **E** to **H** through the right angle between them, advances in the direction of propagation. As the waves corresponding to the

two states of polarization are independent, we shall limit our subsequent discussion to one of them, taking

$$E_y = g(x - \qquad (132\text{–}6)$$

If the wave is simple harmonic g is a sine or cosine function and we can write

$$E_y = A \sin (\alpha x - \omega t), \qquad \alpha \equiv \frac{\omega}{v}. \qquad (132\text{–}7)$$

Denoting the wave-length by λ and the period by P,

$$\alpha(x + \lambda) = \alpha x + 2\pi, \qquad \omega(t + P) = \omega t + 2\pi.$$

Consequently

$$\alpha = \frac{2\pi}{\lambda}, \qquad \omega = \frac{2\pi}{P},$$

and we can write (132–7) in the form

$$E_y = A \sin 2\pi \left(\frac{x}{\lambda} - \frac{t}{P}\right). \qquad (132\text{–}9)$$

Often it is convenient in dealing with plane waves to make use of the *wave-slowness* S which is defined as the reciprocal of the phase velocity v. Then $\alpha = \omega S$ and (132–9) may be written

$$E_y = A \sin \omega(Sx - t). \qquad (132\text{–}10)$$

Finally the last equation may be replaced by

$$E_y = Ae^{i\omega(Sx-t)}, \qquad (132\text{–}11)$$

where it is understood that the imaginary part of the complex expression on the right is to be taken as representing E_y.

If the wave normal or ray does not have the direction of the X axis, we can choose a new set of axes $X'Y'Z'$ (Fig. 296) such that X' is perpen-

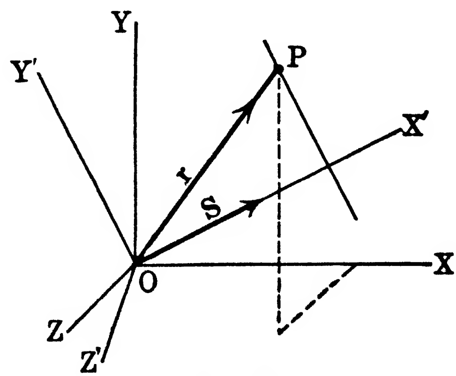


FIG. 296

dicular to the wave front and Y' is parallel to **E**. Then (132–11) holds if we replace x by x' and y by y'. But if x, y, z are the coordinates of a point P in the wave front relative to XYZ,

$$x' = lx + my + nz,$$

where l, m, n are the direction cosines of OX' with respect to the original axes. Consequently

$$E_{y'} = Ae^{i\omega(S[lx+my+nz]-t)} = Ae^{i\omega(\mathbf{S}\cdot\mathbf{r}-t)},$$

where **r** is the position vector of P and **S** has been given the direction of the wave normal. The exponent in the last expression has the advantage of being in vector form and therefore independent of the orientation of the axes to which the wave motion is referred.

Equation (132–4) shows that electromagnetic waves travel *in vacuo* with the velocity c of light. The inference that light itself is electromagnetic in character is inescapable. In addition to being transverse, electromagnetic waves possess many other characteristics of light waves. It is important to note that the wave equations (132–1) and (132–2) would not have followed from the field equations without Maxwell's revision of Ampère's law. It was his introduction of the displacement current which made possible the electromagnetic theory of light.

The index of refraction n of a material medium is defined as the ratio of the velocity of light *in vacuo* to that in the medium. Hence, from (132–4),

$$n = \frac{c}{v} = \sqrt{\kappa\mu}.$$

Both κ and μ may be shown to be functions of the frequency, but for frequencies at all comparable with those of the visible spectrum μ is very exactly unity for all media. In this case we may write

$$n = \sqrt{\kappa}.$$

For infra-red radiation κ should have nearly the limiting value which it assumes for steady fields. The following table illustrates the excellent agreement found for a number of gases.

	κ for Steady Fields	n^2 for Infra-Red Radiation
Air	1.000586	1.000585
H_2	1.000264	1.000277
CO_2	1.000984	1.000909
CO	1.000694	1.000670

The gamut of electromagnetic radiation with which we are familiar today extends all the way from radio waves on the long wave-length end to cosmic rays on the short wave-length end of the spectrum. The different regions comprise roughly the following ranges of wave-length:

Waves from Oscillatory Circuits...............$(10)^4$ km to $(10)^{-2}$ cm,
Infra-Red Radiation........................$(10)^{-1}$ cm to $7(10)^{-5}$ cm,
Visible Spectrum..........................$7(10)^{-5}$ cm to $4(10)^{-5}$ cm,
Ultra-Violet Radiation.....................$4(10)^{-5}$ cm to $(10)^{-6}$ cm,
X-Rays..................................$(10)^{-6}$ cm to $(10)^{-9}$ cm,
γ-Rays..................................$(10)^{-8}$ cm to $(10)^{-10}$ cm,
Cosmic Rays.............................$(10)^{-11}$ cm to $(10)^{-12}$ cm.

All electromagnetic waves travel with the same velocity c in empty space, being distinguished from one another only by differences in wave-length and frequency.

In the next article we shall see how to compute both the energy associated with an electromagnetic field and the flow of energy across a fixed surface through which electromagnetic waves are passing.

Problem 132a. Standing waves are formed by two sets of progressive waves of the same amplitude and wave-length traveling in opposite directions. If the progressive waves are represented by

$$E_y' = A \cos \frac{2\pi}{\lambda} (x - \qquad\qquad E_y'' = A \cos \frac{2\pi}{\lambda} (x + ct),$$

find the resultant E and H and the positions of nodes and loops. The medium is empty space.

$$\text{Ans. } E_y = 2A \cos \frac{2\pi}{\lambda} x \cos \frac{2\pi c}{\lambda} t, \qquad H_z = \quad \sin \frac{2\pi}{\lambda} x \sin \frac{2\pi c}{\lambda} t.$$

133. Electromagnetic Energy and Poynting Flux. — To obtain the energy equation we multiply the three equations (3′) by E_x, E_y, E_z respectively and the three equations (4′) by H_x, H_y, H_z and subtract the sum of the last three from that of the first three. If we pick out the terms containing derivatives with respect to x on the left-hand side of the resulting equation we have

$$-E_y\frac{\partial H_z}{\partial x} + E_z\frac{\partial H_y}{\partial x} - \left(-H_y\frac{\partial E_z}{\partial x} + H_z\frac{\partial E_y}{\partial x}\right)$$
$$= -\frac{\partial}{\partial x}(E_yH_z - E_zH_y) \equiv -\frac{1}{c}\frac{\partial\sigma_x}{\partial x},$$

if we put $\boldsymbol{\sigma}$ for the vector $c(\mathbf{E}\times\mathbf{H})$. So the entire equation is

$$-\left(\frac{\partial\sigma_x}{\partial x} + \frac{\partial\sigma_y}{\partial y} + \frac{\partial\sigma_z}{dz}\right) = \mathbf{E}\cdot\dot{\mathbf{D}} + \mathbf{H}\cdot\dot{\mathbf{B}} + \rho\mathbf{v}\cdot\mathbf{E}.$$

Multiplying by an element $d\tau$ of volume, replacing $\mathbf{D}$ by $\kappa\mathbf{E}$ and $\mathbf{B}$ by $\mu\mathbf{H}$ and rearranging terms, we get

$$\rho\mathbf{v}\cdot\mathbf{E}d\tau + \left(\frac{\partial\sigma_x}{\partial x} + \frac{\partial\sigma_y}{\partial y} + \frac{\partial\sigma_z}{\partial z}\right)d\tau = -\frac{\partial}{\partial t}\left(\tfrac{1}{2}\kappa E^2 + \tfrac{1}{2}\mu H^2\right)d\tau, \tag{133-1}$$

in an isotropic medium of constant permeability.

Consider the flux of $\boldsymbol{\sigma}$ out of a volume element such as that illustrated in Fig. 294. The net flux through the two surfaces perpendicular to the X axis is

$$-\sigma_x dydz + \left(\sigma_x + \frac{\partial\sigma_x}{\partial x}dx\right)dydz = \frac{\partial\sigma_x}{\partial x}d\tau,$$

where $d\tau = dxdydz$. Adding similar expressions for the other two pairs of faces,

$$\text{Outward flux of } \boldsymbol{\sigma} = \left(\frac{\partial\sigma_x}{\partial x} + \frac{\partial\sigma_y}{\partial y} + \frac{\partial\sigma_z}{\partial z}\right)d\tau.$$

The second term on the left-hand side of (133-1), then, represents the flux of the vector $\boldsymbol{\sigma}$ out of the volume $d\tau$. If, now, we integrate the whole equation over an arbitrary volume τ, this term

will specify the flux of $\boldsymbol{\sigma}$ out of the entire volume. This flux, however, can be expressed by the surface integral

$$\int_s \boldsymbol{\sigma} \cdot d\mathbf{s}$$

taken over the closed surface s bounding τ. We have, then, for any volume τ

$$\int_\tau \rho\mathbf{v} \cdot \mathbf{E} d\tau + \int_s \boldsymbol{\sigma} \cdot d\mathbf{s} = -\frac{\partial}{\partial t}\int_\tau (\tfrac{1}{2}\kappa E^2 + \tfrac{1}{2}\mu H^2) d\tau. \qquad (133\text{–}2)$$

This equation is known as the *energy equation*. We shall now investigate the meaning of the three terms which it contains.

The rate at which work is done by the electromagnetic field on the charge contained in τ is obtained by taking the scalar product of the force per unit volume (5), article 130, by $\mathbf{v}$ and integrating over τ. But the vector $\mathbf{v} \times \mathbf{B}$ in the second term of (5) is perpendicular to $\mathbf{v}$ and hence $(\mathbf{v} \times \mathbf{B}) \cdot \mathbf{v} = 0$. The magnetic field, therefore, does no work on the moving charges. The rate of doing work, then, is

$$\int_\tau \mathbf{F}_\tau \cdot \mathbf{v} d\tau = \int_\tau \rho\mathbf{v} \cdot \mathbf{E} d\tau.$$

This is precisely the first term on the left of (133–2).

To interpret the term on the right of (133–2) we shall consider the case where $\mathbf{E}$ and $\mathbf{H}$ vanish at all points on the boundary of τ. Then $\boldsymbol{\sigma}$ is zero over the surface s and (133–2) reduces to

$$\int_\tau \rho\mathbf{v} \cdot \mathbf{E} d\tau = -\frac{\partial}{\partial t}\int_\tau (\tfrac{1}{2}\kappa E^2 + \tfrac{1}{2}\mu H^2) d\tau. \qquad (133\text{–}3)$$

As the work done on the moving charges is performed at the expense of the energy of the field, the right-hand side of this equation must represent the time rate of decrease of electromagnetic energy. Therefore we may represent the energy of the field in the volume τ by the integral

$$U = \int_\tau (\tfrac{1}{2}\kappa E^2 + \tfrac{1}{2}\mu H^2) d\tau, \qquad (133\text{–}4)$$

and ascribe the electric energy

$$u_E = \tfrac{1}{2}\kappa E^2 = \frac{\kappa}{8\pi} E_s^2 \tag{133-5}$$

and the magnetic energy

$$u_H = \tfrac{1}{2}\mu H^2 = \frac{\mu}{8\pi} H_m^2 \tag{133-6}$$

to each unit volume (the subscripts s and m specifying e.s.u. and e.m.u. as usual). These expressions are identical with the electrostatic energy (21–9) and the magnetostatic energy (46–3) per unit volume which we have already found it convenient to attribute to the medium through which the field extends.

We are now ready to consider the complete equation (133–2). In the previous special case there was no possibility of a flow of energy through the surface s bounding the volume τ since the field was zero at all points of this surface. In the general case, however, the law of conservation of energy reads: the rate at which the electromagnetic field does work on the charges in the volume τ plus the rate at which energy flows out in the form of radiation through the surface s bounding τ must equal the rate of decrease of electromagnetic energy in this region. Therefore we conclude that the second term on the left of (133–2) represents the energy passing through the surface s per unit time in the form of radiation. From the form of this term it is clear that we can ascribe a flow of energy of the amount

$$\boldsymbol{\sigma} = c(\mathbf{E} \times \mathbf{H}) = \frac{1}{4\pi}(\mathbf{E}_m \times \mathbf{H}_m) \tag{133-7}$$

per second to each square centimeter of cross-section. This vector is known as the *Poynting flux*. Its direction is that of the flow of energy and its magnitude specifies the quantity of energy passing through a unit cross-section in a unit time. It is to be noticed that $\boldsymbol{\sigma}$ vanishes unless both electric and magnetic fields are present and not in the same direction. The flux of energy is perpendicular to the plane of $\mathbf{E}$ and $\mathbf{H}$ in the sense of advance of a right-handed screw rotated from the first to the second of the

two vectors. It is proportional in magnitude to E, H and the sine of the angle between them.

In the case of the plane wave specified by (132–6) the electric and magnetic energies per unit volume are

$$\left.\begin{aligned} u_E &= \frac{\kappa}{2} E_y^2 = \frac{\kappa}{2} g^2, \\ u_H &= \frac{\mu}{2} H_z^2 = \frac{\kappa}{2} g^2. \end{aligned}\right\} \qquad (133\text{–}8)$$

The energy of the wave, then, is half electric and half magnetic. The Poynting flux is

$$\sigma_x = cE_yH_z = v\kappa g^2 = v(u_E + u_H). \qquad (133\text{–}9)$$

As this represents the energy passing through a unit cross-section in a unit time, the entire energy of the wave progresses with the phase velocity v. This result is independent of direction and we have in general for plane waves

$$\sigma = vu, \qquad (133\text{–}10)$$

where u is the total energy density.

Problem 133a. Calculate the energy of a spherical conductor of radius a charged with a quantity of electricity Q from (133–5) and show that the result agrees with (21–1).

Problem 133b. Calculate the energy of the circuit of Fig. 135 from (133–6) and show that the result agrees with the calculation made in problem 92*a*.

Problem 133c. The earth has a downward electric field of 100 volt/meter and a northward magnetic field of 0.5 gauss. What is the Poynting flux? Ans. 0.40 joule/cm² sec eastward.

Problem 133d. Find the mean energy density $\bar{u}$ and the mean flux of energy $\bar{\sigma}$ for the standing waves of problem 132*a*. Ans. $\bar{u} = \bar{u}' + \bar{u}''$, $\bar{\sigma} = 0$.

Problem 133e. A straight wire of resistance R carries a current i. Compute the Poynting flux through the surface and show that it accounts for the production of heat at the rate Ri^2.

134. Reflection of Electromagnetic Waves by a Perfect Conductor. — Let MN (Fig. 297) be the plane surface of a perfect

conductor on which plane electromagnetic waves are incident at the angle ϕ. As the electric intensity must be zero at every instant in every part of the conductor, it is clear that the wave cannot penetrate into its interior. Therefore we have to consider only the incident radiation BO and the reflected radiation OC.

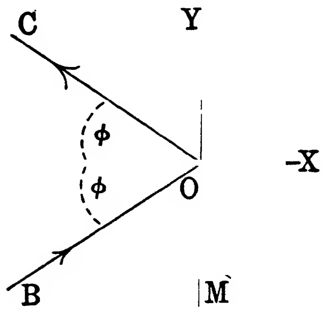


FIG. 297

If we orient axes so that the X axis is perpendicular to MN and the Z axis to the plane of incidence BOC, the vector electric intensity in the incident wave is given by the imaginary part of

$$\mathbf{E} = \mathbf{A}e^{i\omega(S[x\cos\phi + y\sin\phi] - t)}$$

and that in the reflected wave by the imaginary part of

$$\mathbf{E}' = \mathbf{A}'e^{i\omega(S[-x\cos\phi + y\sin\phi] - t)} \qquad (134\text{–}2)$$

from (132–12), where S is the wave-slowness in the medium to the left of MN.

At the surface of the conductor ($x = 0$) the Y and Z components of the resultant electric intensity must vanish. Hence, as the exponentials in the expressions for the two waves become identical for $x = 0$,

$$A_y + A_y' = 0,$$

$$A_z + A_z' = 0.$$

Moreover, as the electric intensity is perpendicular to the direction of propagation,

$$A_x \cos\phi + A_y \sin\phi = 0,$$

$$-A_x' \cos\phi + A_y' \sin\phi = 0,$$

and, as $A_y' = - \quad {}_y$, it follows that $A_x' = A_x$. So we have in all

$$(134\text{–}3)$$

As the amplitude of the reflected wave has the same magnitude as that of the incident wave, all the energy is reflected. Consequently *a perfect conductor is a perfect reflector of electromagnetic waves.*

The resultant electric intensity normal to the surface of the conductor is

$$E_x = 2A_x \sin \omega(Sy \sin \phi - t),$$

and if κ is the dielectric constant of the medium to the left of MN,

$$D_x = 2\kappa A_x \sin \omega(Sy \sin \phi - t).$$

In Heaviside-Lorentz units the flux of D through a closed surface is equal to the charge inside, in accord with (108–3). Therefore the charge per unit area on the surface of the conductor is

$$\sigma = D_x = 2\kappa A_x \sin \omega(Sy \sin \phi - t). \qquad (134\text{–}4)$$

If $A_{\|}$ is the component of the amplitude of the incident radiation in the plane of incidence, $A_x = A_{\|} \sin \phi$ and

$$\sigma = 2\kappa A_{\|} \sin \phi \sin \omega\ (Sy \sin \phi - t), \qquad (134\text{–}5)$$

showing that the surface charge is greatest for grazing incidence and zero for normal incidence. The charge fluctuates from positive to negative and back again during the course of a period.

Although waves incident on a perfect conductor do not pass through the surface, there is always some penetration beyond the surface of actual conductors.

Problem 134a. Show that when electromagnetic waves are reflected at the surface of a perfect conductor the resultant H at the surface is parallel to the surface.

135. Reflection and Refraction of Plane Waves at the Surface of a Dielectric. — When light or electromagnetic waves of any wave-length strike the interface between two dielectrics, the radiation is partly reflected and partly transmitted. We shall consider separately the case where the incident radiation is

polarized with the electric vector perpendicular to the plane of incidence and the case where the electric vector lies in the plane of incidence. Let MN (Fig. 298) be the surface separating a medium of specific inductive capacity κ_1 on the left from a medium of specific inductive capacity κ_2 on the right. The incident, reflected and refracted rays are BO, OC and OD respectively, ϕ_1 being the angle of incidence and ϕ_2 the angle of refraction. Our aim is to find the ratio of the amplitude A_1' of the reflected radiation to the amplitude A_1 of the incident radiation, and of the amplitude A_2 of the transmitted radiation to the amplitude A_1 of the incident radiation. The first of these ratios is known as the *coefficient of reflection* R and the second as the *coefficient of transmission* T. Since the energy density is proportional to the square of A, the ratio of the intensity of the reflected to that of the incident radiation, known as the *reflecting power*, is equal to the square of the coefficient of reflection. Incidentally we shall deduce Snell's law giving the ratio of sin ϕ_2 to sin ϕ_1.

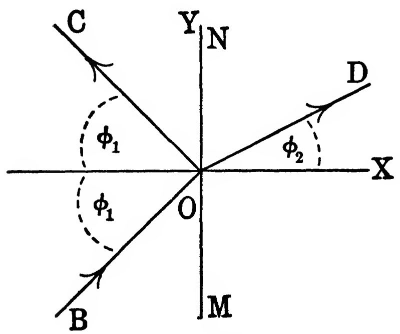


FIG. 298

At the surface MN the normal component of **D** and the tangential component of **E** must be continuous, as proved in article 14. Since μ is always unity for high frequency fields, all components of **H** must be continuous. These boundary conditions have to be satisfied, then, by the three systems of waves.

Perpendicular Case. — If **E** is perpendicular to the plane of incidence, we have only the component E_z. We write then

$$\left.\begin{aligned} \mathbf{E}_1 &= kA_1 e^{\iota\omega(S_1[x\cos\phi_1 + y\sin\phi_1] - t)}, \\ \mathbf{E}_1' &= kA_1' e^{\iota\omega(S_1[-x\cos\phi_1 + y\sin\phi_1] - t)}, \\ \mathbf{E}_2 &= kA_2 e^{\iota\omega(S_2[x\cos\phi_2 + y\sin\phi_2] - t)}, \end{aligned}\right\} \quad (135\text{–}1)$$

for the incident, reflected and refracted wave trains respectively,

S_1 being the wave-slowness in the first medium and S_2 that in the second. In each of these expressions the electric intensity is represented by the imaginary part of the complex quantity.

The magnetic vector is perpendicular to the electric vector in the sense which makes $\mathbf{E} \times \mathbf{H}$ have the direction of propagation. So, from (132–6),

$$\left.\begin{aligned} \mathbf{H}_1 &= n_1 E_1(\mathbf{i} \sin \phi_1 - \mathbf{j} \cos \phi_1), \\ \mathbf{H}_1' &= n_1 E_1'(\mathbf{i} \sin \phi_1 + \mathbf{j} \cos \phi_1), \\ \mathbf{H}_2 &= n_2 E_2(\mathbf{i} \sin \phi_2 - \mathbf{j} \cos \phi_2), \end{aligned}\right\} \quad (135\text{–}2)$$

where we have replaced $\sqrt{\kappa}$ by the index of refraction n in accord with (132–14).

The boundary conditions, therefore, are

$$E_1 + E_1' = E_2, \qquad (135\text{–}3)$$

$$(E_1 + E_1')n_1 \sin \phi_1 = E_2 n_2 \sin \phi_2, \qquad (135\text{–}4)$$

$$(E_1 - E_1')n_1 \cos \phi_1 = E_2 n_2 \cos \phi_2, \qquad (135\text{–}5)$$

for $x = 0$, the first expressing the continuity of the tangential component of $\mathbf{E}$ and the remaining two the continuity of the two components of $\mathbf{H}$. Dividing (135–4) by (135–3) we have *Snell's law*,

$$\frac{\sin \phi_1}{\sin \phi_2} = \frac{n_2}{n_1} = \frac{v_1}{v_2} = \frac{S_2}{S_1}. \qquad (135\text{–}6)$$

As a consequence of this relation the exponentials in (135–1) become identical for $x = 0$. Hence we may replace the E's in (135–3), (135–4) and (135–5) by the A's. Dividing (135–5) by (135–4) to eliminate A_2, and putting $R_\perp = A_1'/A_1$,

$$\frac{1 - R_\perp}{1 + R_\perp} = \frac{\sin \phi_1 \cos \phi_2}{\sin \phi_2 \cos \phi_1},$$

which gives for the coefficient of reflection

$$R_\perp = \frac{\sin \phi_2 \cos \phi_1 - \sin \phi_1 \cos \phi_2}{\sin \phi_2 \cos \phi_1 + \sin \phi_1 \cos \phi_2} = -\frac{\sin (\phi_1 - \phi_2)}{\sin (\phi_1 + \phi_2)}. \qquad (135\text{–}7)$$

Similarly, if we solve for $T_\perp = A_2/A_1$, we get for the coef-

ficient of transmission

$$T_{\perp} = \frac{2 \sin \phi_2 \cos \phi_1}{\sin (\phi_1 + \phi_2)}. \qquad (135\text{–}8)$$

Parallel Case. — If the electric vector lies in the plane of incidence the electric intensities in the incident, reflected and refracted waves are respectively

$$\left.\begin{aligned} \mathbf{E}_1 &= A_1(-\boldsymbol{i} \sin \phi_1 + \boldsymbol{j} \cos \phi_1)e^{i\omega(S_1[x \cos \phi_1 + y \sin \phi_1] - t)}, \\ \mathbf{E}_1' &= A_1'(-\boldsymbol{i} \sin \phi_1 - \boldsymbol{j} \cos \phi_1)e^{i\omega(S_1[-x \cos \phi_1 + y \sin \phi_1] - t)}, \\ \mathbf{E}_2 &= A_2(-\boldsymbol{i} \sin \phi_2 + \boldsymbol{j} \cos \phi_2)e^{i\omega(S_2[x \cos \phi_2 + y \sin \phi_2] - t)}, \end{aligned}\right\} \qquad (135\text{–}9)$$

and the magnetic intensity is

$$\left.\begin{aligned} \mathbf{H}_1 &= \boldsymbol{k} n_1 E_1, \\ \mathbf{H}_1' &= \boldsymbol{k} n_1 E_1', \\ \mathbf{H}_2 &= \boldsymbol{k} n_2 E_2. \end{aligned}\right\} \qquad (135\text{–}10)$$

Equating the normal components of **D** and the tangential components of **E** and **H** on the two sides of the surface of separation we have the boundary conditions

$$(E_1 + E_1')n_1^2 \sin \phi_1 = E_2 n_2^2 \sin \phi_2, \qquad (135\text{–}11)$$

$$(E_1 - E_1') \cos \phi_1 = E_2 \cos \phi_2, \qquad (135\text{–}12)$$

$$(E_1 + E_1')n_1 = E_2 n_2, \qquad (135\text{–}13)$$

if we replace κ by n^2. The quotient of (135–11) by (135–13) leads to Snell's law again. Solving for the coefficient of reflection $R_{\parallel} = A_1'/A_1$ and the coefficient of transmission $T_{\parallel} = A_2/A_1$ we find in this case

$$R_{\parallel} = \frac{\sin \phi_1 \cos \phi_1 - \sin \phi_2 \cos \phi_2}{\sin \phi_1 \cos \phi_1 + \sin \phi_2 \cos \phi_2} = \frac{\tan (\phi_1 - \phi_2)}{\tan (\phi_1 + \phi_2)}, \qquad (135\text{–}14)$$

and

$$T_{\parallel} = \frac{2 \sin \phi_2 \cos \phi_1}{\sin (\phi_1 + \phi_2) \cos (\phi_1 - \phi_2)}. \qquad (135\text{–}15)$$

The theoretical expressions for the coefficients of reflection and of transmission for both states of polarization are well verified

by experiment. In fact, Fresnel had given these relations as empirical formulas long before the advent of the electromagnetic theory of light.

It is interesting to note that of the four coefficients only $R_{\|}$ can vanish. This occurs when $\phi_1 + \phi_2 = \pi/2$, for then the denominator $\tan(\phi_1 + \phi_2)$ becomes infinite. Under these circumstances the reflected and refracted rays are at right angles, and all incident radiation polarized with the electric vector parallel to the plane of incidence is transmitted. If the incident radiation contains both states of polarization, the reflected radiation consists solely of electromagnetic waves polarized with the electric vector perpendicular to the plane of incidence. So by allowing unpolarized light to strike the surface of a dielectric such as glass at the proper angle a reflected beam of polarized light may be obtained. However, on account of surface irregularities, the polarization is never complete.

The angle of incidence Φ_1 necessary to make $R_{\|}$ vanish is known as *Brewster's angle* or the *polarizing angle.* Making use of Snell's law we have

$$n_1 \sin \Phi_1 = n_2 \sin \left(\frac{\pi}{2} - \Phi_1 \right),$$

or

$$\Phi_1 = \text{arc tan} \left(\frac{n_2}{n_1} \right). \tag{135-16}$$

For glass in air $n_2/n_1 = 1.5$ and consequently $\Phi_1 = 56°.3$.

Problem 135a. Show that the reflecting power of glass of index of refraction n for normal incidence is $\left(\frac{n-1}{n+1}\right)^2$. What change in phase takes place on reflection? Why do $R_\perp$ and $R_{\|}$ differ in sign for normal incidence?

136. Radiation Pressure and Electromagnetic Momentum. — When electromagnetic waves impinge on matter they exert a stress known as *radiation pressure.* To calculate this stress we must express the force per unit volume (5), article 130, in terms

of the components of **E** and **H** alone. Writing down the X component of $\mathbf{F}_\tau$ we have

$$F_{\tau_x} = \rho E_x + \frac{1}{c}(\rho v_y B_z - \rho v_z B_y). \qquad (136\text{–}1)$$

The charge ρ can be eliminated by means of (1′), and the components of the current density $\rho\mathbf{v}$ by means of (3′), giving

$$F_{\tau_x} = E_x\left(\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}\right) + B_z\left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} - \frac{1}{c}\dot{D}_y\right)$$
$$- B_y\left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - \frac{1}{c}\dot{D}_z\right). \qquad (136\text{–}2)$$

Furthermore if we multiply (2′) by H_x and the second and third equations of (4′) by D_z and D_y respectively and combine,

$$0 = H_x\left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z}\right) + D_z\left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} + \frac{1}{c}\dot{B}_y\right)$$
$$- D_y\left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} + \frac{1}{c}\dot{B}_z\right). \qquad (136\text{–}3)$$

Adding (136–2) and (136–3), putting $D = \kappa E$ and $B = \mu H$ and collecting terms,

$$F_{\tau_x} = \frac{\partial}{\partial x}\left\{\frac{\kappa}{2}(E_x^2 - E_y^2 - E_z^2) + \frac{\mu}{2}(H_x^2 - H_y^2 - H_z^2)\right\}$$
$$+ \frac{\partial}{\partial y}\{\kappa E_x E_y + \mu H_x H_y\} + \frac{\partial}{\partial z}\{\kappa E_x E_z + \mu H_x H_z\}$$
$$- \frac{1}{c}\frac{\partial}{\partial t}\{\kappa\mu(E_y H_z - E_z H_y)\}. \qquad (136\text{–}4)$$

Consider a small rectangular parallelopiped (Fig. 299) of matter of dimensions Δx, Δy, Δz. If we designate the stresses in the X direction on its faces by X_x, X_y, X_z, the first being a tension on the face perpendicular to the X axis and the second and third shears on the faces perpendicular respectively to the Y and Z axes, the total force in the X direction is

$$\left(\frac{\partial X_x}{\partial x} + \frac{\partial X_y}{\partial y} + \frac{\partial X_z}{\partial z}\right)\Delta x \Delta y \Delta z,$$

or

$$\frac{\partial X_x}{\partial x} + \frac{\partial X_y}{\partial y} + \frac{\partial X_z}{\partial z}$$

per unit volume. Comparing with (136–4) we see that the elec-

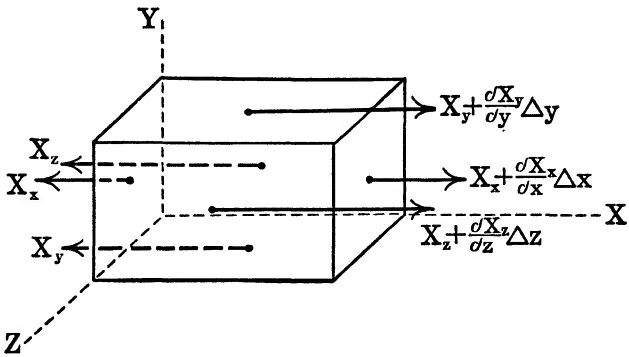


FIG. 299

trical stresses parallel to the X axis are

$$\left.\begin{aligned} X_x &= \frac{\kappa}{2}(E_x^2 - E_y^2 - E_z^2) + \frac{\mu}{2}(H_x^2 - H_y^2 - H_z^2), \\ X_y &= \kappa E_x E_y + \mu H_x H_y, \\ X_z &= \kappa E_x E_z + \mu H_x H_z. \end{aligned}\right\} \quad (136\text{–}5)$$

In addition to this system of stresses the electromagnetic field exerts a body force the X component of which is given by the last term in (136–4). Now $(E_yH_z - E_zH_y)$ is the X component of the vector $\mathbf{E} \times \mathbf{H}$. Consequently the body force is

$$\mathbf{F}_\tau' = -\frac{\partial}{\partial t}\left(\frac{\kappa\mu}{c}\mathbf{E}\times\mathbf{H}\right) = -\frac{\partial}{\partial t}\left(\frac{\kappa\mu}{c^2}\boldsymbol{\sigma}\right) \quad (136\text{–}6)$$

per unit volume, where $\boldsymbol{\sigma}$ is the Poynting flux $c(\mathbf{E} \times \mathbf{H})$.

Consider a slab of matter MN (Fig. 300) on the left-hand face of which a steady train of electromagnetic waves impinges normally. As $\mathbf{E}$ and $\mathbf{H}$ are perpendicular to the direction of propagation, $E_x = H_x = 0$. Hence the tension is

$$X_x = -\left(\tfrac{1}{2}\kappa E^2 + \tfrac{1}{2}\mu H^2\right) = -u, \quad (136\text{–}7)$$

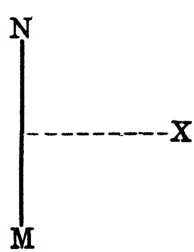


FIG. 300

where u is the energy density of the radiation. As the stress is negative it represents a pressure.

If $\bar{u}_1$ is the mean energy density in the incident radiation, $\bar{u}_1'$ that in the reflected radiation and $\bar{u}_2$ that in the radiation which passes through the slab, the average force per unit area of the slab in the X direction due to the radiation pressure under consideration is

$$\bar{p} = \bar{u}_1 + \bar{u}_1' - \bar{u}_2. \tag{136–8}$$

In the case of a perfect reflector, this becomes $\bar{p} = 2\bar{u}_1$.

The shear in the Y direction is

$$Y_x = \kappa E_y E_x + \mu H_y H_x,$$

obtained by interchanging x and y in the second of equations (136–5). This vanishes since $E_x = H_x = 0$. The same is true of Z_x. Finally the average value of the body force $\mathbf{F}_\tau'$ on the slab vanishes, since the intensity of the incident radiation remains constant and therefore the mean value of σ does not alter with the time. Consequently the total force is given by (136–8). The predicted pressure has been verified experimentally by Nichols and Hull and by Lebedew.

Consider next a region τ containing radiation but no matter. The force $\mathbf{F}_\tau$, then, must be zero. If $\mathbf{K}$ is the force on this region due to the stresses acting over its surface,

$$0 = \mathbf{K} - \frac{\partial}{\partial t}\int_\tau \left(\frac{\kappa\mu}{c^2}\sigma\right) d\tau.$$

As the volume τ contains radiation and nothing else, the force $\mathbf{K}$ due to the stresses on its surface must be supposed to act on the radiation in its interior. As we have

$$\mathbf{K} = \frac{\partial}{\partial t}\int_\tau \left(\frac{\kappa\mu}{c^2}\sigma\right) d\tau, \tag{136–9}$$

and as force is equal to time rate of increase of momentum, the radiation has an *electromagnetic momentum*

$$\mathbf{G} = \int_\tau \left(\frac{\kappa\mu}{c^2}\sigma\right) d\tau. \tag{136–10}$$

We may therefore attribute momentum in the amount

$$\mathbf{g} = \frac{\kappa\mu}{c^2}\boldsymbol{\sigma} \qquad (136\text{–}11)$$

to each unit volume of a radiation field. For plane waves the electromagnetic momentum per unit volume has the magnitude

$$g = \frac{\kappa\mu}{c^2}vu = \frac{u}{v} \qquad (136\text{–}12)$$

from (133–10) and (132–4), that is, equal to the energy per unit volume divided by the velocity of propagation.

To illustrate the utility of the concept of electromagnetic momentum we shall calculate the pressure on the perfectly reflecting surface *MN* (Fig. 297) due to a train of electromagnetic waves incident at the angle ϕ. If v is the velocity of propagation and g_1 the momentum per unit volume of the incident radiation, a momentum $g_1 v \cos\phi$ is transferred from the incident to the reflected beam per unit time by each unit area of the reflector. Consequently the time rate of increase of electromagnetic momentum of the radiation field is $2g_1 v \cos^2\phi$ to the left. This is the force exerted by each unit area of *MN* on the radiation. According to the law of action and reaction the radiation field exerts an equal and opposite force on the reflecting surface. The latter therefore is subject to the pressure $2g_1 v \cos^2\phi$. Now if $\bar{u}_1$ is the mean energy density of the incident radiation, the mean pressure is

$$\bar{p} = 2\bar{g}_1 v \cos^2\phi = 2\bar{u}_1 \cos^2\phi \qquad (136\text{–}13)$$

from (136–12). Note that equations (136–8) and (136–13) give the same result for normal incidence on a perfect reflector.

Problem 136a. The earth receives 20 large calories (1 large calorie = 1000 calories) per square meter per minute from the sun. Calculate the maximum electric intensity and magnetic intensity in sunlight. Assuming that all the light is absorbed, compute the mean pressure at normal incidence. Ans. $E = 10.2$ volt/cm, $H = 0.0341$ gauss, $\bar{p} = 4.6(10)^{-5}$ dyne/cm^2.

137. Electromagnetic Oscillators. — While we have considered in some detail the properties of plane electromagnetic waves, we have given no attention to possible sources of electromagnetic radiation. In all cases electromagnetic waves originate in the oscillation of electric charge. The necessary vibratory motion may be produced most simply by means of the Hertzian oscillator described in article 89. Here an oscillatory discharge takes place across the gap G (Fig. 192), the charges on A and B reversing sign every half period. Essentially we have equal positive and negative charges vibrating along the line AB with simple harmonic motion, the charge of the one sign being out of phase with that of the other by π radians. The mechanism is equivalent to an oscillating electric doublet. In our theoretical investigation of the production of electromagnetic waves we shall suppose that the oscillations are maintained at constant amplitude so as to avoid the complications produced by damping and shall assume that the region surrounding the oscillator is empty space.

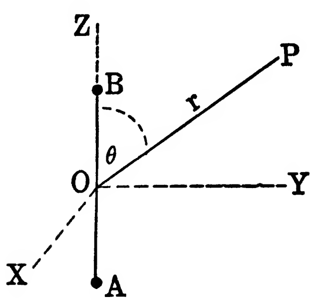


FIG. 301

Oscillating Electric Doublet. — Consider a positive charge q oscillating with frequency $\omega/2\pi$ along the section AB (Fig. 301) of the Z axis. If a_1 is the amplitude of vibration, the displacement of the charge at any instant is

$$z_1 = a_1 \cos \omega t.$$

The displacement of an equal negative charge vibrating with the same frequency but out of phase by π radians with the first is

$$z_2 = a_2 \cos (\omega t + \pi) = - a_2 \cos \omega t,$$

where a_2 represents the amplitude of the second charge. At any instant, then, the distance between the charges is

$$z = z_1 - z_2 = (a_1 + a_2) \cos \omega t,$$

and the electric moment of the doublet is

$$p = qz = p_0 \cos \omega t, \qquad p_0 \equiv q(a_1 + a_2). \qquad (137\text{–}1)$$

Close to the doublet the field at any instant may be treated as if static, the components of the electric intensity being given by (12–2). Consequently, in Heaviside-Lorentz units,

$$\left.\begin{aligned} E_r &= \frac{2p_0 \cos\theta}{4\pi r^3}\cos\omega t, \\ E_\theta &= \frac{p_0 \sin\theta}{4\pi r^3}\cos\omega t. \end{aligned}\right\} \tag{137–2}$$

We get the magnetic field near the doublet from (69–2). After transforming to Heaviside-Lorentz units,

$$\begin{aligned} H &= \frac{q}{4\pi c r^2}\left(\frac{dz_1}{dt} - \frac{dz_2}{dt}\right)\sin\theta \\ &= -\frac{\omega p_0 \sin\theta}{4\pi c r^2}\sin\omega t \end{aligned} \tag{137–3}$$

at right angles to the rZ plane. The lines of magnetic force are therefore circles about the Z axis.

As its electromagnetic field is propagated outwards from the oscillator with a finite velocity, equations (137–2) and (137–3), in which we have treated the field as if it were static, are valid only in the neighborhood of the source. To obtain $\mathbf{E}$ at a distance from the oscillator we must solve the wave equation (132–1) subject to Gauss' law (1″). The solution must reduce to (137–2) in the immediate vicinity of the oscillator if the waves are to be those produced by the mechanism which we have postulated. In fact we may consider (137–2) to represent the *boundary conditions* which our solution must satisfy.

Since the wave equation (132–1) is linear with constant coefficients any derivative with respect to x, y, z or t of a solution is also a solution. We shall start, then, with a solution which is a function of r and t alone, and build up other solutions by differentiation. As

$$r^2 = x^2 + y^2 + z^2,$$

it follows that

$$\frac{\partial}{\partial x} = \frac{\partial r}{\partial x}\frac{\partial}{\partial r} = \frac{x}{r}\frac{\partial}{\partial r},$$

$$\frac{\partial^2}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{x}{r}\frac{\partial}{\partial r}\right) = \frac{x^2}{r^2}\frac{\partial^2}{\partial r^2} - \frac{x^2}{r^3}\frac{\partial}{\partial r} + \frac{1}{r}\frac{\partial}{\partial r},$$

and similar expressions hold for $\frac{\partial^2}{\partial y^2}$ and $\frac{\partial^2}{\partial z^2}$. Consequently

$$\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}=\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}.$$

So if we take for the dependent variable in (132–1) a functior Φ_0 of r and t alone, the wave equation becomes

$$\frac{\partial^2\Phi_0}{\partial r^2}+\frac{2}{r}\frac{\partial\Phi_0}{\partial r}=\frac{1}{c^2}\frac{\partial^2\Phi_0}{\partial t^2},$$

since $\kappa=\mu=1$. As all the derivatives are partial, this can b(written

$$\frac{\partial^2}{\partial r^2}(r\Phi_0)=\frac{1}{c^2}\frac{\partial^2}{\partial t^2}(r\Phi_0), \qquad (137\text{–}4)$$

an equation which is of precisely the same form as (132–3) Consequently its solution is

$$\Phi_0=\frac{1}{r}f_1(r-ct)+\frac{1}{r}f_2(r+ct). \qquad (137\text{–}5)$$

Evidently the first term represents a spherical wave divergin from the origin and the second term one converging toward th(origin. We are interested in a simple harmonic diverging wave Therefore we shall take

$$\Phi_0=\frac{A}{r}\cos(\alpha r-\omega t), \qquad \frac{\omega}{\alpha}=c, \qquad (137\text{–}6)$$

in the terminology of (132–7).

Since Φ_0 is a solution of the wave equation,

$$\Phi_1\equiv-\frac{\partial\Phi_0}{\partial z}=-\frac{\partial\Phi_0}{\partial r}\cos\theta$$

$$=A\cos\theta\left\{\frac{1}{r^2}\cos(\alpha r-\omega t)+\frac{\alpha}{r}\sin(\alpha r-\omega t)\right\} \qquad (137\text{–}7)$$

is also. If, then, we take for the components of $\mathbf{E}$

$$E_x=-\frac{\partial\Phi_1}{\partial x},\quad E_y=-\frac{\partial\Phi_1}{\partial y},\quad E_z=-\frac{\partial\Phi_1}{\partial z}-\frac{1}{c^2}\frac{\partial^2\Phi_0}{\partial t^2}, \qquad (137\text{–}8)$$

each component of the electric intensity is a solution of the wave equation.

As this procedure seems somewhat artificial at first sight, it may be illuminating to observe that the method we are following is closely analogous to that which we might have pursued in obtaining the components (12–2) of the electric intensity due to a static dipole. In that case we might start with the solution

$$V_0 = \frac{A}{r}$$

of Laplace's equation, then differentiate with respect to z to get the potential (12–1)

$$V_1 = -\frac{\partial V_0}{\partial z} = \frac{A \cos\theta}{r^2}$$

of the dipole, and finally write

$$E_x = -\frac{\partial V_1}{\partial x}, \qquad E_y = -\frac{\partial V_1}{\partial y}, \qquad E_z = -\frac{\partial V_1}{\partial z},$$

a set of equations which, of course, is exactly equivalent to (12–2). The only respects in which our procedure in obtaining (137–8) differs from that which we have outlined above for the static dipole are that we start with a solution of the wave equation, which is the generalization of Laplace's equation for electromagnetic fields, and that we have an added term in the expression (137–8) for E_z. The necessity of this added term will appear immediately.

We have yet to show that the components (137–8) of the electric intensity satisfy Gauss' law (1″). Making use of (137–7),

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = \frac{\partial}{\partial z}\left\{ \frac{\partial^2 \Phi_0}{\partial x^2} + \frac{\partial^2 \Phi_0}{\partial y^2} + \frac{\partial^2 \Phi_0}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 \Phi_0}{\partial t^2} \right\}.$$

But, as Φ_0 is a solution of the wave equation, the expression within the braces vanishes identically. Consequently (1″) is satisfied. It should be noted that if we had omitted the second term in the expression for E_z, the components of **E**, while they would have satisfied the wave equation, would have failed to satisfy Gauss' law.

Now equation (137–8) shows that E_x, E_y and the first term in E_z are obtained from Φ_1 just as the components of the electric intensity (7–4) are obtained from the potential in an electrostatic field. So we can consider that we have the field that would be produced by the potential Φ_1 plus the electric intensity $-\frac{1}{c^2}\frac{\partial^2\Phi_0}{\partial t^2}$ in the direction of the axis of the doublet. Consequently we have from (7–5)

$$\left.\begin{aligned} E_r &= -\frac{\partial\Phi_1}{\partial r} - \frac{1}{c^2}\frac{\partial^2\Phi_0}{\partial t^2}\cos\theta, \\ E_\theta &= -\frac{\partial\Phi_1}{r\partial\theta} + \frac{1}{c^2}\frac{\partial^2\Phi_0}{\partial t^2}\sin\theta, \end{aligned}\right\} \qquad (137\text{–}9)$$

for the components of electric intensity in the directions of increasing r and θ respectively. Making use of (137–6) and (137–7), we get explicitly

$$\left.\begin{aligned} E_r &= A\cos\theta\left\{\frac{2}{r^3}\cos(\alpha r-\omega t) + \frac{2\alpha}{r^2}\sin(\alpha r-\omega t)\right\}, \\ E_\theta &= A\sin\theta\left\{\left(\frac{1}{r^3}-\frac{\alpha^2}{r}\right)\cos(\alpha r-\omega t)\right. \\ &\qquad\qquad\left. + \frac{\alpha}{r^2}\sin(\alpha r-\omega t)\right\}. \end{aligned}\right\} \qquad (137\text{–}10)$$

As $\alpha = 2\pi/\lambda$, the terms in $1/r^3$ predominate over the others at distances from the origin small compared to the wave-length. Furthermore, near the origin αr in the argument of the trigonometrical functions represents a negligible phase angle. Consequently we have

$$\left.\begin{aligned} E_r &= \frac{2A\cos\theta}{r^3}\cos\omega t, \\ E_\theta &= \frac{A\sin\theta}{r^3}\cos\omega t, \end{aligned}\right\} \qquad (137\text{–}11)$$

in the immediate neighborhood of the oscillating doublet. So the boundary conditions (137–2) are satisfied if the constant A is made equal to the amplitude p_0 of the electric moment divided by 4π.

At a great distance from the origin we need retain only those terms in (137–10) involving $1/r$. Hence E_r has a negligible magnitude compared to E_θ, and, putting $2\pi/\lambda$ for α and c for ω/α, the electric intensity consists of the component

$$E_\theta = -\frac{\pi p_0 \sin\theta}{\lambda^2 r}\cos\frac{2\pi}{\lambda}(r - ct) \qquad (137\text{–}12)$$

alone. At this distance from the doublet a small section of the wave front is approximately plane, and the magnetic intensity is given by (132–6). It is, then,

$$H_\phi = -\frac{\pi p_0 \sin\theta}{\lambda^2 r}\cos\frac{2\pi}{\lambda}(r - ct) \qquad (137\text{–}13)$$

in a direction at right angles to the rZ plane, that is, in the direction of increasing azimuth ϕ measured about the Z axis. Therefore the lines of magnetic force are circles around the Z axis.

Both E_θ and H_ϕ are zero along the axis of the dipole on account of the factor $\sin\theta$ and are maximum in the equatorial plane. The radiation, therefore, is emitted from the dipole in greatest intensity in directions at right angles to the line of oscillation. Both fields fall off inversely with the distance r and, for a constant p_0, are inversely proportional to the square of the wavelength or directly proportional to the square of the frequency. Inspection of (137–10) shows that the phase of E_θ changes by π as we pass from the vicinity of the origin to a distant point. While the two fields are in phase at a great distance from the source, comparison of (137–2) and (137–3) shows that they differ in phase by $\pi/2$ in the neighborhood of the oscillating doublet. That the latter must be the case is clear from the fact that when the electric moment is greatest the current is zero and *vice versa*. The radiation emitted is polarized with E in the rZ plane and H at right angles thereto.

The Poynting flux is

$$\sigma = cE_\theta H_\phi = \frac{\pi^2 c p_0^2 \sin^2\theta}{\lambda^4 r^2}\cos^2\frac{2\pi}{\lambda}(r - ct) \qquad (137\text{–}14)$$

in the direction of the radius vector.

The energy $\mathcal{P}_r$ radiated per unit time is obtained by integrating σ over the surface of a sphere of radius r. We have then

$$\mathcal{P}_r = \int_0^\pi \sigma \cdot 2\pi r^2 \sin\theta d\theta = \frac{8\pi^3 c p_0^2}{3\lambda^4} \cos^2 \frac{2\pi}{\lambda}(r - ct), \quad (137\text{–}15)$$

and the mean rate of radiation is

$$\overline{\mathcal{P}}_r = \frac{4\pi^3 c p_0^2}{3\lambda^4}. \quad (137\text{–}16)$$

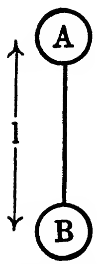


Fig. 302

For a given p_0 the rate of radiation is inversely proportional to the fourth power of the wave-length or directly proportional to the fourth power of the frequency.

In an actual oscillatory circuit the charges which act as sources of electromagnetic radiation are not point charges as in the ideal oscillator which we have discussed. Rather we have a continuous back and forth flow of electricity between the conductors A and B (Fig. 302). As

$$p = ql = p_0 \cos \omega t,$$

the current i is

$$i = \frac{dq}{dt} = i_0 \sin \omega t,$$

where

$$i_0 l = -\omega p_0 = -\frac{2\pi c}{\lambda} p_0.$$

In terms of the current amplitude (137–12) and (137–13) become

$$E_\theta = H_\phi = \frac{i_0 l \sin\theta}{2\lambda c r} \cos\frac{2\pi}{\lambda}(r - ct), \quad (137\text{–}17)$$

and the mean rate of radiation of energy is

$$\overline{\mathcal{P}}_r = \frac{\pi i_0^2 l^2}{3\lambda^2 c} = \frac{4\pi^2 c (i_0)_m^2 l^2}{3\lambda^2}, \quad (137\text{–}18)$$

where the subscript m indicates e.m.u. as usual.

If $(R_0)_m$ is the ohmic resistance of the oscillatory circuit, the

rate at which energy is converted into heat is

$$\overline{\mathcal{P}}_0 = \tfrac{1}{2}(R_0)_m(i_0)_m^2$$

from (90–8). So the total dissipation of energy per unit time is

$$\overline{\mathcal{P}} = \overline{\mathcal{P}}_0 + \overline{\mathcal{P}}_r = \frac{1}{2}\left\{ (R_0)_m + \frac{8\pi^2 c l^2}{3\lambda^2} \right\} (i_0)_m^2.$$

The effective resistance of the circuit, then, is increased by radiation in the amount

$$(R_r)_m = \frac{8\pi^2 c l^2}{3\lambda^2} \qquad (137\text{–}19)$$

in e.m.u. This is known as the *radiation resistance* of the oscillator. In practical units (ohms),

$$(R_r)_p = \frac{80\pi^2 l^2}{\lambda^2}. \qquad (137\text{–}20)$$

The radiation characteristics of the ordinary vertical wire antenna used in wireless communication can be determined with the aid of (137–12) and (137–13). The antenna is divided into a number of short lengths each of which acts like a Hertzian oscillator. The total radiation field is then the sum of the fields due to these elementary oscillators, and the radiation resistance of the antenna can be calculated by the same methods that we have used here.

Oscillating Magnetic Doublet. — A circular loop of wire C (Fig. 303) in which an oscillating current $i = i_0 \sin \omega t$ is flowing constitutes an oscillating magnetic doublet. For, on account of the back and forth flow of current, the field which it produces is the same as that of an equivalent magnetic shell whose polarity is a simple harmonic function of the time. If A is the area of the circuit the equivalent magnetic moment M is the product of A by the current in e.m.u., or

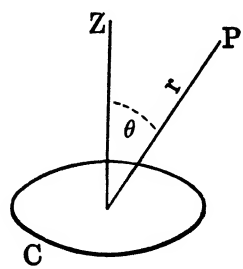


FIG. 303

$$M = \frac{iA}{c}$$

in h.l.u., and the amplitude of the magnetic moment is

$$M_0 = \frac{i_0 A}{c}.$$

All we need do to obtain the radiation field is to interchange E and H in (137–12) and (137–13) and replace p_0 by M_0. We have then

$$E_\phi = H_\theta = -\frac{\pi i_0 A \sin\theta}{\lambda^2 cr}\cos\frac{2\pi}{\lambda}(r - ct). \qquad (137\text{–}21)$$

The mean rate of radiation is

$$\bar{\mathcal{P}}_r = \frac{4\pi^3 i_0{}^2 A^2}{3\lambda^4 c} \qquad (137\text{–}22)$$

from (137–16). Comparing with (137–18) we see that for the same current and wave-length

$$\frac{\text{Radiation from Magnetic Oscillator}}{\text{Radiation from Electric Oscillator}} = \left(\frac{2\pi A}{\lambda l}\right)^2. \qquad (137\text{–}23)$$

Since the linear dimensions of an oscillator must of necessity be small compared to the wave-length λ of the emitted radiation, a magnetic oscillator is inefficient compared with an electric oscillator as a source of electromagnetic waves. Now open circuits operate effectively as Hertzian or electric oscillators, whereas closed loops have the characteristics of magnetic oscillators. Consequently the former are used in preference to the latter for wireless antennas.

Problem 137a. An oscillator has a length of 5 meters and carries a current whose r.m.s. value is 5 amperes. The frequency is 1000 kilocycles. Find the power radiated. Ans. 5.48 watt.

Problem 137b. What is the radiation resistance of the oscillator of 137*a*? Ans. 0.219 ohm.

Problem 137c. Find the relative efficiency as sources of electromagnetic waves of an electric oscillator of length 5 meters and a magnetic oscillator of the same diameter at a frequency of 1000 kilocycles. Ans. 148 : 1.

138. Graphical Representation of Electromagnetic Waves. — Since electromagnetic waves progress through empty space with the velocity c, we can consider the field of an oscillating point charge as traveling out from it radially in all directions with the speed of light, the charge replenishing the field as fast as it moves away by the emission of new lines of force. For instance, let us consider a charge oscillating up and down along the line AC (Fig. 304) with simple harmonic motion and let us fix our attention on a line of electric force extending to the right. If we

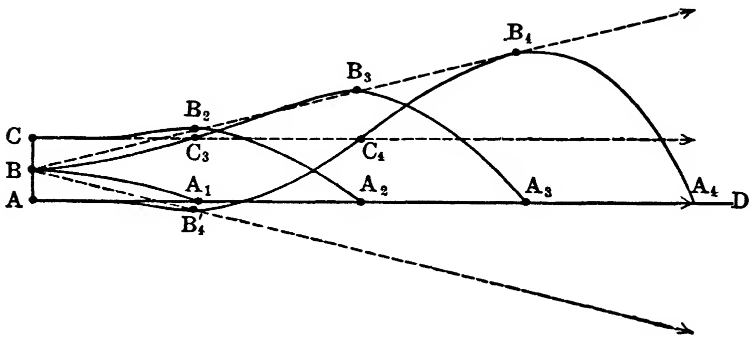


FIG. 304

suppose the charge to be initially at rest at A, the line of force under consideration is then the straight line AD. If, now, the simple harmonic motion is imparted to the charge, the field at A moves out to A_1 with the velocity c while the charge moves up to B. The line of force is now BA_1D. Allowing another equal interval of time to elapse, A_1 has progressed to A_2 and B to B_2 while the charge has reached the upper end C of its path, and the line of force under consideration has assumed the form CB_2A_2D. The direction of motion BB_2 of the point B of the field is horizontally to the right *relative to an observer moving with the charge*, therefore inclined upward relative to an observer who does not partake of the motion. Continuing the process we find $BC_3B_3A_3D$ and $AB_4'C_4B_4A_4D$ for successive later configurations of the line of force under consideration, the line of motion BB_4' being inclined below the horizontal on account of the downward motion of the

charge on passing through B on its return to A. We have here a moving picture of an electromagnetic wave progressing to the right from the oscillating charge. The farther the wave travels from the source, the steeper and more nearly transverse it becomes. The reader can easily show that no wave is emitted in the vertical direction. This is in accord with the analytical result (137–12) obtained in the last article.

The line of force shown in the figure has exactly the configuration of a stream of water issuing from a nozzle kept pointed to the right and caused to oscillate up and down along the line AC. The picture is not merely approximately correct, but is an *exact* representation of an electromagnetic wave. In fact the entire group of Maxwell's field equations, (1″) to (4″) of article 132, has been shown* to be merely the kinematical equations of motion of the lines of force of the field as here represented. Had this representation been available before Oersted's discovery of the magnetic effect of a current, Ampère's law with Maxwell's addition of the displacement current, Faraday's law and the force equation (5) could have been predicted without any appeal to experiment.

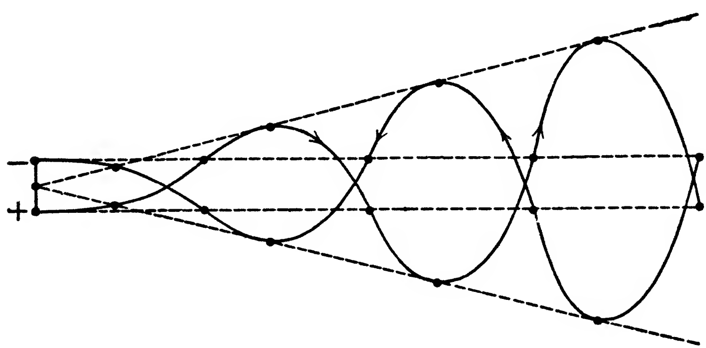


Fig. 305

We cannot correctly apply the foregoing representation to the resultant field of an oscillating doublet since the radial velocity of propagation c applies only to the individual field of a single

* Page, Introduction to Electrodynamics.

point charge. The waves produced by a doublet must rather be considered as the resultant of the two trains of waves to which the two charges give rise. These are illustrated in Fig. 305 for a horizontal line of force extending from each charge. Remembering that the line of force terminating on the positive charge is directed *away* from the charge on which it originates whereas that terminating on the negative charge is directed *toward* its source, it is clear from the diagram that the transverse components of the two fields reinforce each other whereas the longitudinal components are oppositely directed. While the field is largely longitudinal close to the doublet, it is nearly altogether transverse at a distance of a few wave-lengths.

CHAPTER XVII

HIGH FREQUENCY OSCILLATIONS

139. Characteristics of High Frequency Circuits. — Although high frequency circuits are subject to the same fundamental laws as low frequency circuits, their behavior in practice differs in a number of important respects. In the first place, it is very difficult to construct circuit elements which are pure resistances, inductances or capacities and which have constant magnitudes. Thus a resistance unit usually has a small but not negligible reactance. Moreover, the magnitude of the resistance itself varies somewhat with frequency due to change in the current distribution within the conductors of which the resistance element is composed, an effect discussed in more detail in the next article. Similarly an inductance coil has a resistance which increases appreciably with frequency. Also a coil behaves like an inductance shunted by a small capacity, because of the capacity between adjacent parts of its winding. Again a condenser has an apparent series resistance due to dielectric losses. This is true even of an air condenser, unless it is very carefully constructed.

Distributive effects constitute another difference between high frequency and low frequency circuits. That is, the inductance of leads and connecting wires and the capacity between different parts of a circuit usually cannot be neglected. Stray coupling and stray capacities to nearby objects are particularly troublesome, it often being necessary to resort to shielding as well as to special arrangement of the circuit to avoid errors due to these causes.

Measurements in high frequency circuits canot be made by means of bridges and similar arrangements suitable for low frequency measurements, both because of the lack of constancy of circuit elements and because of the distributive effects just mentioned. Measurements are almost always made to depend

on resonance phenomena, since resonance adjustments are relatively simple at high frequencies. Incidentally, knowledge of the frequency usually is required, so that use of a wave-meter (art. 125) is involved.

Finally, with the exception of thermocouple meters (art. 64) instruments suitable for low frequency circuits cannot be used in high frequency circuits. Special devices are required, the most important of which is the thermionic vacuum tube described in article 141.

High frequency theory and practice contain so many distinctive features that they are often regarded as constituting a separate field of physical knowledge. Some of the most important of these features are discussed in this chapter.

140. Resistance, Inductance and Capacity. — As has been pointed out in the previous article circuit elements do not possess pure resistance, inductance or capacity of constant magnitude in high frequency circuits. Let us investigate the high frequency characteristics of resistance units, inductance coils and condensers.

Resistance. — Consider first a straight conductor of circular cross-section carrying a current $i = i_0 \sin \omega t$. Let the length of the conductor be l and its radius a. We are interested in the current density j at different points in the conductor. At low frequencies the current density is everywhere $j = j_0 \sin \omega t$, where $j_0 = i_0/\pi a^2$, since the current is uniformly distributed. At high frequencies, however, the current density is not uniform across the cross-section of the conductor, for the change of magnetic flux inside the conductor results in an e.m.f. of sufficient magnitude to affect the distribution of current. It is evident from symmetry that the internal e.m.f. and consequently the current density are functions of the distance from the axis of the conductor.

To find the current density we may use the electromagnetic field equations developed in Chapter XVI. However, the solution of these equations under the given conditions is rather

difficult. As the variation of the current density is quite small, even at frequencies of several million cycles, we can obtain a satisfactory approximate solution of the problem by supposing that the total current density J at any point is made up of the low frequency density j and a circulating or eddy current density j_e. That is,

$$J = j + j_e. \qquad (140\text{–}1)$$

The eddy currents, which are due to the internal e.m.f. in the conductor, do not affect the total amount of current passing, of course, but merely alter its distribution. In calculating j_e we shall take account of the flux corresponding to the undisturbed current density j, but shall neglect the small amount of flux due to j_e itself.

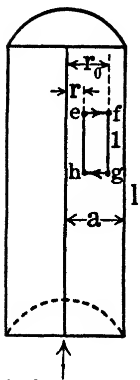


Fig. 306

The eddy current density is a function of the distance r (Fig. 306) from the axis of the conductor, so that the eddy currents flow in cylindrical shells. Moreover, as the eddy currents do not produce any resultant flow of current, there must be some distance r_0 from the axis at which j_e is zero, the eddy currents always being in opposite directions inside and outside of r_0. Referring to the figure, let us determine the e.m.f. around the rectangle $efgh$ of unit height lying in a plane through the axis. With the positive direction of current in the conductor as indicated, the positive direction of e.m.f. around the rectangle is in the sense of the arrows.

If we exclude ferromagnetic materials and assume l large compared to a, the flux of induction in e.m.u. through the rectangle is

$$\int_r^{r_0} \frac{2i_0 \sin \omega t}{a^2} r dr = \pi(r_0^2 - r^2) j_0 \sin \omega t,$$

using (72–6). Then, from Faraday's law (85–1),

$$\frac{j_e}{\sigma} = -\frac{d}{dt}[\pi(r_0^2 - r^2) j_0 \sin \omega t],$$

where σ is the conductivity of the conductor, so that

$$j_e = \pi\sigma\omega(r^2 - r_0^2)j_0 \cos \omega t.$$

It remains to determine r_0. As the total flow of current due to j_e is always zero we have

$$\int_0^a j_e \cdot 2\pi r dr = 0,$$

or, removing factors independent of r,

$$\int_0^a (r^2 - r_0^2) r dr = \frac{a^4}{4} - \frac{r_0^2 a^2}{2} = 0.$$

Thus $r_0 = a/\sqrt{2}$ and, finally,

$$j_e = \pi\sigma\omega\left(r^2 - \frac{a^2}{2}\right)j_0 \cos \omega t. \qquad (140\text{–}2)$$

Evidently the mean rate at which energy is dissipated in the conductor for a given magnitude i_0 of total current flowing is greater at high frequencies than at low due to the presence of the eddy currents. The low frequency rate is $\frac{1}{2}Ri_0^2$ from (90–8), where $R = l/\pi a^2\sigma$ is the low frequency or d.c. resistance. In order to express the high frequency rate in the same form it is convenient to define the *effective resistance* R_e by setting $\frac{1}{2}R_e i_0^2$ equal to the mean power developed in the conductor. As the instantaneous power is

$$\int_0^a \frac{J^2 l}{\sigma} 2\pi r dr = \frac{j_0^2 l}{\sigma}\int_0^a \left[\sin^2 \omega t + 2\pi\sigma\omega\left(r^2 - \frac{a^2}{2}\right)\sin \omega t \cos \omega t + \pi^2\sigma^2\omega^2\left(r^2 - \frac{a^2}{2}\right)^2 \cos^2 \omega t\right] 2\pi r dr,$$

the mean power is

$$\frac{1}{2}\frac{j_0^2 l}{\sigma}\int_0^a \left[1 + \pi^2\sigma^2\omega^2\left(r^2 - \frac{a^2}{2}\right)^2\right] 2\pi r dr = \frac{1}{2}Ri_0^2\left[1 + \frac{2\pi^2\sigma^2\omega^2}{a^2}\int_0^a \left(r^2 - \frac{a^2}{2}\right)^2 r dr\right],$$

since in taking mean values $\overline{\sin^2 \omega t} = \overline{\cos^2 \omega t} = \frac{1}{2}$ while

$\overline{\sin \omega t \cos \omega t} = 0$. Thus, equating $\frac{1}{2}R_e i_0^2$ to the expression just obtained and evaluating the integral,

$$\frac{1}{2} R_e i_0^2 = \frac{1}{2} R i_0^2 \left[1 + \frac{2\pi^2\sigma^2\omega^2}{a^2} \frac{a^6}{24} \right],$$

or, in terms of the frequency ν,

$$\frac{R_e}{R} = 1 + \frac{\pi^4\sigma^2\nu^2 a^4}{3}. \qquad (140\text{–}3)$$

In applying this formula it must be remembered that σ is in e.m.u. although R_e and R may be expressed in any system of units since they appear as a ratio.

The increase of the effective resistance of a conductor with frequency is called the *skin effect*, because the current density diminishes along the axis and increases near the surface. This fact does not appear from our approximate analysis which is valid only for small values of $\sigma\nu a^2$, but is brought out by an exact solution of the problem.

To minimize the skin effect we must keep $\sigma\nu a^2$ small. The sizes of manganin and copper wire (American wire gauge) which must be used to keep R_e/R less than 1.01 are shown for various frequencies in the table. Evidently as far as skin effect goes it is desirable to use fine wires of low conductivity. For this reason

ν	Wire Size (A. W. G.)	
	Manganin	Copper
$1.0(10)^5$	14	28
2.5	18	31
5.0	21	34
7.5	23	36
10.0	24	38

primary resistance standards for high frequencies usually consist of single lengths of fine manganin wire. For convenience these resistance units are made up in sets covering a range of resistance values. The construction of such a set is described in *Circular 74 of the Bureau of Standards.*

When resistance wire is wound into the form of a coil the variation of the effective resistance with frequency is much increased, for now there are eddy current losses in each turn of wire due not only to its own field, the skin effect, but also to the field of adjacent turns, the *proximity effect.* In addition there is always some residual inductance even with the non-inductive type of winding and some capacity between different parts of the coil, so that at high frequencies its reactance is appreciable. Resistance boxes can be used therefore only when great accuracy is not required and then only subject to frequency limitations. For general purposes carefully constructed resistance boxes may be used up to 50,000 cycles without incurring errors greater than five or ten per cent. In tuned circuits where the tuning neutralizes the residual reactance of resistance elements the frequency limit is considerably higher.

Inductance. — An inductance coil is subject to two disturbing influences at high frequencies. The skin and proximity effects cause the effective resistance R_e to increase quite rapidly with frequency, while the capacity between different parts of the winding causes the current to vary along the coil. In the case of a single layer coil, for example, the current is greater in the middle than at the ends. It is found that the impedance of the

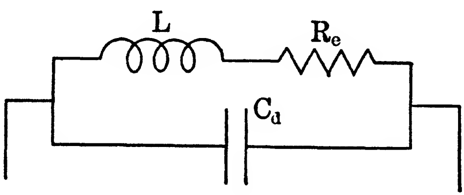


FIG. 307

coil is approximately equivalent to that of the divided circuit shown in Fig. 307, L being the low frequency inductance and C_d, by definition, the *distributed capacity.* Referring to article 112, the complex impedance of the divided circuit is seen to be

$$\frac{R_e}{(1 - LC_d\omega^2)^2 + R_e^2C_d^2\omega^2} + i\frac{L\omega(1 - LC_d\omega^2) - R_e^2C_d\omega}{(1 - LC_d\omega^2)^2 + R_e^2C_d^2\omega^2}.$$

As C_d is of the order of a few micromicrofarads and the coil is never used with a frequency so high that $LC_d\omega^2$ approaches unity we can neglect $R_eC_d\omega$ in comparison with $(1 - LC_d\omega^2)$ and the expression for the impedance reduces to

$$\frac{R_e}{(1 - LC_d\omega^2)^2} + i\frac{L\omega}{(1 - LC_d\omega^2)}.$$

Thus the *apparent resistance* R_a and the *apparent inductance* L_a of the coil are given by

$$\left.\begin{aligned} R_a &= \frac{R_e}{(1 - LC_d\omega^2)^2}, \\ L_a &= \frac{L}{(1 - LC_d\omega^2)}. \end{aligned}\right\} \quad (140\text{–}4)$$

It appears that both the apparent resistance and the apparent inductance increase with frequency due to the distributed capacity. The former also increases due to the skin and proximity effects, so its rate of increase is relatively large.

As (140–4) gives only approximate results, particularly in the case of the resistance, and as the values of R_a and L_a in practice depend somewhat on the position of the coil in the circuit, stray capacity to ground, the presence of nearby conductors and so on, it is necessary in accurate work to measure R_a and L_a under the exact conditions in which the coil is to be used. Methods of measurement are described in article 142.

Capacity. — The behavior of a condenser at high frequencies is similar to its behavior at low frequencies. There are energy losses due to dielectric absorption (art. 20) whose effect may be represented by an apparent series resistance R_C. If C is the capacity of the condenser its power factor is given by

$$\cos\phi = R_C C\omega,$$

as in (120–11).

Although the power factor does not change greatly with frequency it is usually somewhat greater at high frequencies than at low. Also it is more important at high frequencies, for cir-

cuits are usually adjusted to resonance and small resistance is desirable in order to obtain large currents. For these reasons care must be used in selecting condensers for high frequency circuits. Condensers in which the dielectric is waxed paper have a power factor of the order of $(10)^{-2}$, which is often too large to be satisfactory. Mica condensers, and air condensers in which the insulation between the two sets of plates is not well arranged, have a power factor of about $(10)^{-3}$. In the case of a carefully built air condenser in which the plates are supported by some low loss material such as quartz the power factor may be as small as $5(10)^{-5}$. If the condenser is variable this figure applies at maximum capacity, the power factor being approximately inversely proportional to the capacity setting.

As the high frequency characteristics of standard air condensers can be accurately determined at low frequencies, using methods described in Chapter XIV, high frequency measurements are made as far as possible in terms of capacity.

141. Vacuum Tubes. — The *vacuum tube* is a device whose operation depends primarily on the control of space charge (art. 80). The simplest and one of the most useful types is the three-electrode tube or *triode*. This consists of a hot *filament F* (Fig. 308), a fine wire *grid G* and a *plate P*, all enclosed in a highly evacuated bulb. The tube whose elements are shown in the diagram has a cylindrical structure but the electrodes may equally well be parallel planes. Thus the triode is similar to the two-electrode tube described in article 80 except for the grid which has been interposed between the cathode and the anode.

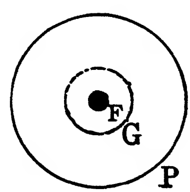


FIG. 308

Let us suppose that potentials $\mathcal{E}_G$ and $\mathcal{E}_P$, measured relative to the filament, are applied to the grid and plate, respectively, the latter potential being much greater than the former. Electrons are drawn from the filament through the grid to the plate. Assuming a sufficiently plentiful supply of electrons, the plate current i_P is determined by the distribution of space charge

between F and P. Now, as we have seen, with no grid present the plate voltage controls the space charge in such a way that i_P is proportional to $\mathcal{E}_P^{3/2}$. With the grid, however, the space charge depends on both $\mathcal{E}_G$ and $\mathcal{E}_P$. It is found that because the grid is much closer than the plate to the filament $\mathcal{E}_G$ is more effective than $\mathcal{E}_P$ in controlling the space charge. In fact, a given change in grid voltage causes the same change in plate current as a change in plate voltage μ times as great, where μ is a constant called the *amplification factor*. This constant ranges in value from two or three to several hundred depending on the geometry of the tube. Thus for the triode the plate current is proportional to $(\mathcal{E}_P + \mu\mathcal{E}_G)^{3/2}$ and we may write

$$i_P = K_P(\mathcal{E}_P + \mu\mathcal{E}_G)^{3/2}, \tag{141–1}$$

K_P being a constant.

The way in which i_P varies with $\mathcal{E}_G$ and $\mathcal{E}_P$ is shown graphically in Fig. 309 for a typical case. Each curve represents the variation of plate current with grid voltage for a definite value of plate voltage. In practice the grid voltage is restricted to negative and small positive values, as large positive values cause electrons to be deflected to the grid diminishing the plate current and reducing the effectiveness of the grid control.

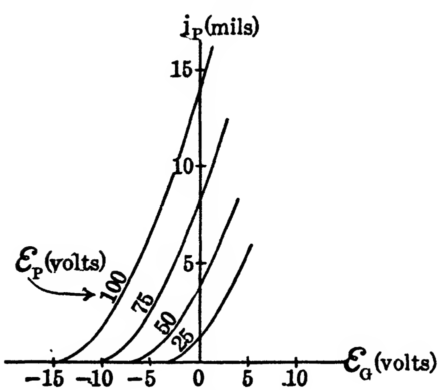


FIG. 309

There are a number of uses to which vacuum tubes may be put. The most important of these in connection with high frequencies are *amplification*, *detection* and *oscillation*, which we shall discuss in turn. In each case small variable e.m.f.'s and currents are superposed on larger steady ones.

Consider the typical circuit arrangement shown in Fig. 310. The tube elements are indicated schematically in the circle. Voltage for the filament, plate voltage and grid voltage are supplied by batteries $\mathcal{E}_A$, $\mathcal{E}_B$ and $\mathcal{E}_C$. The subscripts *A*, *B* and *C* are customarily employed to indicate quantities associated with the filament, plate and grid circuits respectively, and the

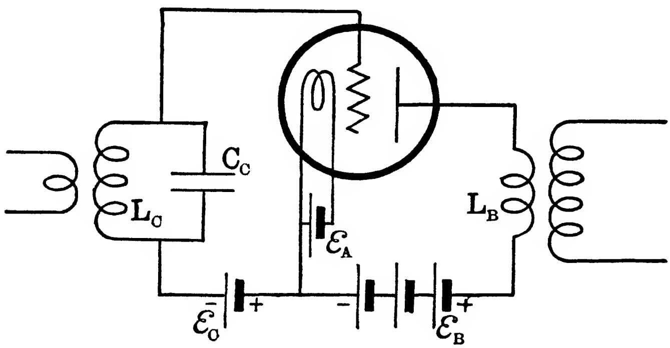


FIG. 310

three sources of e.m.f. just mentioned are often called the *A*-battery, the *B*-battery and the *C*-battery. A high frequency e.m.f. is introduced into the grid circuit by means of two coils inductively coupled, the coil L_C in the grid circuit being tuned to the applied frequency by a condenser C_C. Similarly, a high frequency e.m.f. is taken from the plate circuit by means of inductively coupled coils, the coil L_B in the plate circuit carrying the steady component of the plate current as well as the varying.

Suppose now that $\mathcal{E}_B$ has a relatively large value, while $\mathcal{E}_C$ is negative, so that the characteristic $i_P \sim \mathcal{E}_G$ curve and the *operating point* O are as shown in Fig. 311. Then, as O is on the part of the curve which is nearly straight, the variable part of the plate current has essentially the same form as the variable part of the grid voltage. This is illustrated by the two sinusoidal curves on the diagram. Finally, the voltage taken from the plate circuit also has the same form as the input voltage. Its amplitude, however, is much greater than that of the input

voltage when the circuit is properly adjusted. Thus the tube functions as an *amplifier*. Several amplifying tubes may be connected in series, the output voltage of one tube supplying the

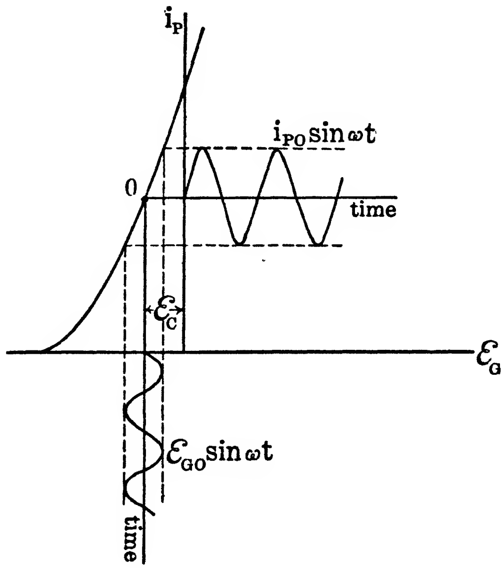


FIG. 311

input voltage to the next. In this way very large amplification may be obtained, an important matter in the reception of the high frequency electromagnetic waves used in wireless telegraphy and telephony.

Next let us change the value of $\mathcal{E}_C$ so as to make O lie on the non-linear part of the $i_P \sim \mathcal{E}_G$ curve (Fig. 312). Then the variable part of the plate current no longer has the same form as the variable part of the grid voltage, but is distorted as illustrated in the diagram. Evidently the mean value of the variable part of the plate current is not zero but has a positive magnitude. Thus, if the plate circuit inductance L_B is replaced by a d.c. milliammeter, the reading will be greater when the high frequency voltage is impressed on the grid than when the grid voltage consists of $\mathcal{E}_C$ alone. The tube is now acting as a *detector* of high frequency

oscillations. As the increase of the meter reading over the steady state value increases with the amplitude of the impressed grid voltage, the meter may be calibrated to indicate this amplitude directly. The vacuum tube and its associated circuit

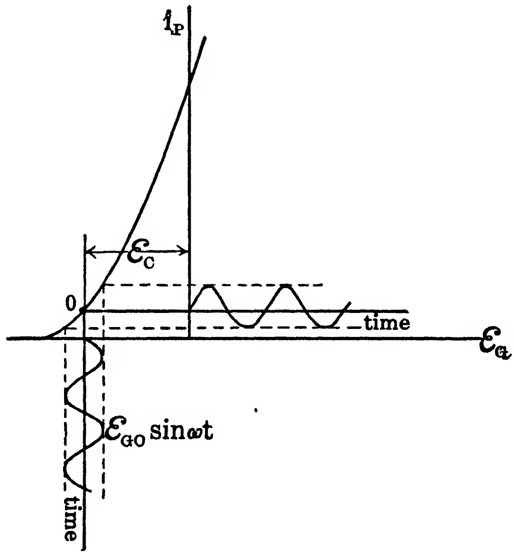


FIG. 312

are then called a *vacuum tube voltmeter.* By using this instrument to measure the e.m.f. across a known impedance the amplitude of a small high frequency current may be determined.

Finally we must consider the production of high frequency oscillations by means of a vacuum tube. If a single tuned circuit is set into oscillation by any means the amplitude of the oscillations diminishes rapidly, as demonstrated in article 89, the energy of the oscillations being dissipated in the resistance of the circuit. If, however, the oscillating circuit is associated with a vacuum tube in such a way that energy is supplied as fast as it is dissipated, then steady oscillations of constant amplitude are maintained. There are a number of ways in which this can be done.

A convenient circuit is shown in Fig. 313. The oscillating

circuit consists of an inductance L_B with resistance R_B shunted by a variable capacity C_B. The oscillating current in L_B induces a voltage in the grid coil L_C to which it is inductively coupled. This voltage in turn causes a plate current which supplies the necessary energy to the oscillating circuit. In order that the phase relations may be correct with the coils connected as shown the windings must be in the same sense. The frequency of oscillation is controlled by the oscillating circuit, being given by $1/2\pi\sqrt{L_B C_B}$ as usual. As the tuned circuit is in the plate circuit of the tube the circuit arrangement of Fig. 313 is called a *tuned-plate oscillator.*

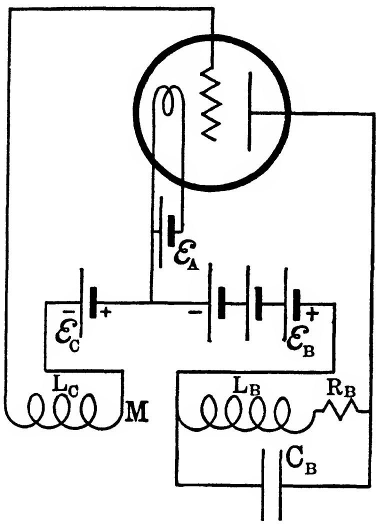


Fig. 313

In order to adjust the oscillator in practice it is usually desirable that the coils be provided with taps as indicated in Fig. 314. Then by moving the grid lead the proper amount of grid voltage

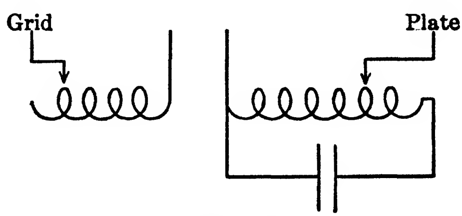


Fig. 314

to sustain oscillation is obtained and by moving the plate lead the amplitude of the oscillating current is controlled.

For a description of other types of oscillators and for a detailed discussion of amplifiers and detectors the reader should refer to some work devoted to the theory of vacuum tubes. No

attempt has been made in this article to do more than give a qualitative description of the uses of vacuum tubes in high frequency circuits.

142. High Frequency Measurements. — In making measurements at high frequencies certain general precautions must be taken in order to obtain consistent and significant results. Thus, a circuit must be arranged so that capacity between its different elements is avoided. Also, the effect of capacity to ground should be minimized by grounding the circuit at the point where this capacity is greatest. Similarly stray inductive coupling must be eliminated. Often shielding is necessary to prevent undesirable coupling, both capacitative and inductive. Finally, in comparison measurements where one circuit element is substituted for another, care must be taken to see that the circuit is not changed in any way by the substitution. The practice of these precautions and some other special ones will be made clear in the description of actual measurements below.

As we are usually interested in determining the complete characteristics of a resistance element, or an inductance coil, or a condenser, we shall discuss the measurements under these headings.

Resistance Element. — The circuit used to calibrate a resistance element at high frequency is shown in Fig. 315. L is an inductance coil loosely coupled to a vacuum tube oscillator, C is a variable condenser and A is a sensitive thermocouple meter.

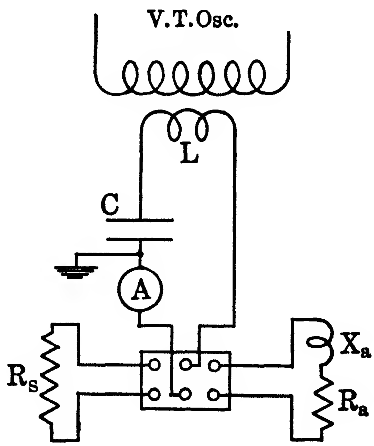


FIG. 315

R_S is a standard variable resistance or a set of standard units such as were described in article 140, while R_a is the apparent resistance of the element to be calibrated and X_a is its apparent

inductance, if any. Either R_S or R_a may be introduced into the circuit by means of a double-pole double-throw switch which may well consist of six small mercury cups and two copper jumpers, instead of the ordinary knife blade type, in order to avoid contact resistance and stray capacity effects. It will be noticed that the leads from L are fairly long in order to remove the rest of the apparatus from the field of the coils. On the other hand, the leads connecting R_S and R_a to the switch are short and of approximately the same length in order that the circuit may not be affected by the substitution of one resistance for the other. One side of the condenser is grounded. If the condenser is enclosed in a metal case this is included in the ground connection.

The actual measurements consist first in adjusting the oscillator to a desired frequency with the aid of a wave-meter (not shown in the figure) as explained in article 126, including R_a in the circuit and tuning by means of C, resonance being indicated by a maximum current reading on the meter. Then R_S is substituted for R_a, the circuit retuned and R_S adjusted until the resonance current has the same value as for R_a. Since the reactance is zero in both cases equal currents indicate equal resistances. Thus $R_a = R_S$. The process is now repeated with a new frequency and so on until the desired frequency range has been covered. It is usually convenient to plot R_a as a function of frequency. If desired, the apparent reactance X_a of the resistance element may be calculated in terms of L and C at each of the test frequencies, since at resonance $L\omega - 1/C\omega + X_a = 0$. Often, however, knowledge of the apparent reactance is not required.

Inductance Coil. — The apparent inductance L_a of a coil, its true inductance L and its distributed capacity C_d can all be obtained from one set of measurements. The experimental arrangement is indicated in Fig. 316, the coil being represented by its equivalent circuit as explained in article 140. C is a calibrated variable condenser and A a thermocouple meter to indicate resonance as in the case of the resistance measurements

described above. The coil is loosely coupled to a vacuum tube oscillator whose frequency can be adjusted with the aid of a wave-meter to any desired value.

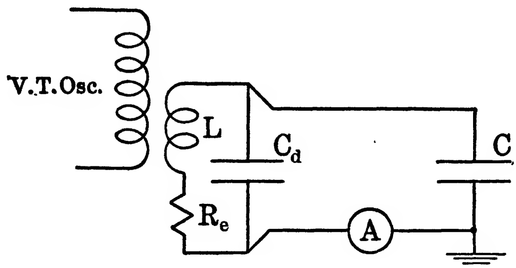


FIG. 316

From (140–4) the apparent inductance at an angular frequency ω is given by

$$L_a = \frac{L}{1 - LC_d\omega^2}.$$

Now, having adjusted the measuring circuit to resonance, we have $1 - L(C + C_d)\omega^2 = 0$. Hence eliminating L we find

$$= \frac{\quad}{C\omega^2}, \tag{142–1}$$

or, eliminating ω,

$$L_a = \tag{142–2}$$

Thus to find L_a as a function of frequency we need only tune the circuit to resonance for a number of frequencies and use (142–1). To find L and C_d we may proceed in either of two ways. Suppose we choose a pair of frequencies ω_1 and ω_2 not too widely separated and tune the circuit for each, C_1 and C_2 being the two resonance values of the capacity. Then

$$1 - L(C_1 + C_d)\omega_1^2 = 0,$$
$$1 - L(C_2 + C_d)\omega_2^2 = 0.$$

Eliminating first C_d and then L we have

$$\left.\begin{aligned} L &= \frac{\dfrac{1}{\omega_1^2} - \dfrac{1}{\omega_2^2}}{C_1 - C_2}, \\ C_d &= \frac{C_1\omega_1^2 - C_2\omega_2^2}{\omega_2^2 - \omega_1^2}. \end{aligned}\right\} \quad (142–$$

In making determinations of L and C_d by this method sever pairs of values in different parts of the frequency range shoul be used as these quantities, particularly C_d, may not be al solutely constant.

The second method is graphical and hence somewhat le accurate than the first. Having tuned the circuit to resonanc for a number of frequencies, we plot $1/\omega^2$ against C (Fig. 317 and draw a straight lin through the points so ol tained, extending the lin until it cuts the C axi Then from the relatio $1/\omega^2 = L(C + C_d)$ we se that the slope of the lin gives L and the intercept o the C axis gives C_d. Whe these quantities have bee determined L_a may readil be calculated with the aid of (142–2) and plotted against fre quency, if it has not been previously obtained.

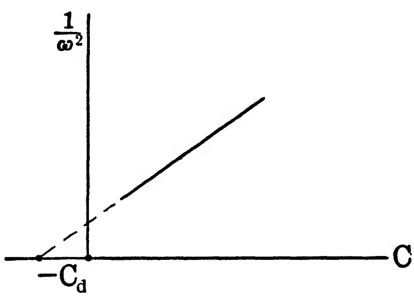


Fig. 317

In applying formulas (142–1) to (142–3) inclusive it must b remembered that the angular frequency ω is 2π times the actua frequency given by the wave-meter.

To determine the apparent resistance R_a of the coil w require a somewhat different circuit, illustrated in Fig. 318. Th coil in question is represented by L_a, R_a while L' is a small coup ling coil loosely coupled to the oscillator. C is a variable con denser, R_S a standard adjustable resistance and A a thermo couple meter. By means of a switch the coil can be introduce

into the circuit or removed from it as desired. The circuit is arranged in accord with the precautions outlined at the beginning of the article, special care being taken that there is no coupling between L_a and the oscillator or L'.

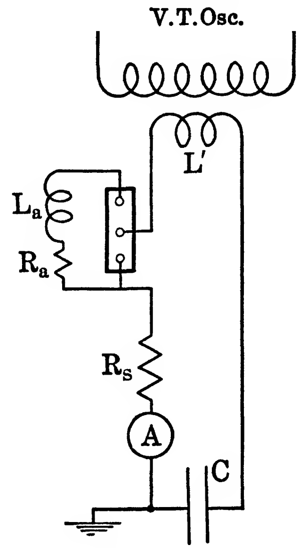


FIG. 318

We first measure the resistance of the entire circuit including R_a. Let R be the resistance of the circuit exclusive of R_a and R_S. Having adjusted the oscillator to a desired frequency we tune the circuit to resonance with the standard resistance set at zero and note the amplitude i_{10} of the current. Then we increase the standard resistance to a value R_S, so that the current amplitude diminishes to i_{20}. This gives

$$\frac{i_{10}}{i_{20}} = \frac{R_a + R + R_S}{R_a + R},$$

or, solving for $R_a + R$,

$$R_a + R = \qquad (142\text{–}4)$$

This result rests on the assumption that there is no change in the e.m.f. induced in L' by the oscillator, that is, that the coupling is loose enough so that the measuring circuit does not react on the oscillator and change the oscillating current. It is important to verify this assumption, either by including a meter in the oscillating circuit of the oscillator or by using several values of R_S and observing whether there is a change in the values obtained for $R_a + R$.

Next we remove the coil, retune the circuit and again measure the resistance. This time we have

$$= R_S \quad \frac{i_{20}'}{' - i_{20}'} \qquad (142\text{–}5)$$

the prime indicating the new set of current and resistance values. R_a is now obtained by taking the difference of (142–4) and (142–5). If the current amplitude is reduced one-half by the addition of the known resistance in each case, we have simply $R_a = R_S - R_S'$. Evidently the condenser must be sufficiently good so that the change in its apparent resistance during the process of retuning is negligible.

When R_a has been determined for a number of frequencies we plot it as a function of frequency, preferably on the same diagram with L_a for the sake of comparison.

Condenser. — The capacity of a condenser is found by comparison with a variable standard condenser. At the same time its apparent resistance R_C may be determined if the standard is of the low loss type so that its resistance is negligible. The circuit arrangement is shown in Fig. 319. The condenser to be calibrated is represented by C, R_C and the standard by C_S. L is a variable but not necessarily calibrated inductance. L' is a small coil loosely coupled to the vacuum tube oscillator. R_S is a standard adjustable resistance and A a thermocouple meter. The circuit is laid out in a manner similar to that of Fig. 315, the condensers being removed

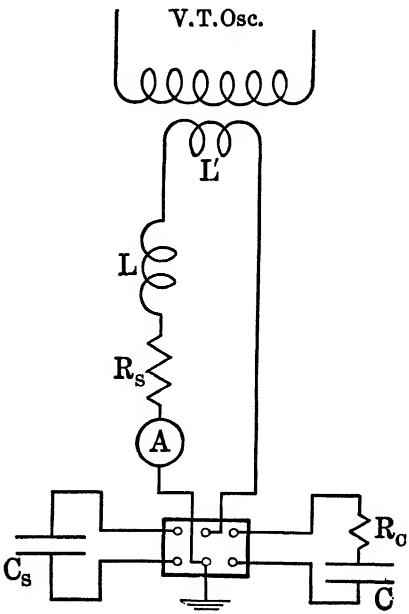


Fig.

from the fields of the coils and symmetrically disposed with respect to the rest of the circuit. One side and the case of each condenser is grounded. The switch by means of which

one condenser is substituted for the other should be of the mercury cup type described in connection with Fig. 315.

To make a determination of C and R_C we adjust the oscillator to a desired frequency with the aid of a wave-meter, include C in the circuit and, with the adjustable resistance set at zero, tune the circuit to resonance by varying L. Then, having noted the amplitude of the resonance current, we replace C by C_S and tune again, varying C_S this time and leaving L unchanged. Finally we set R_S so that the current is the same as before. Under these conditions

$$C = C_S, \qquad R_C = R_S.$$

Unless the condenser is poorly constructed C is independent of frequency and the power factor $R_C C\omega$ is very nearly so. However, it is usually desirable to repeat the measurements at several other frequencies to verify this behavior.

If the condenser is variable we determine C and R_C for a number of settings at one frequency and then repeat for a few settings only at several other frequencies. It is usually convenient to plot C as a function of the condenser reading on a fairly large scale in order to be able to determine the capacity at any setting accurately.

A more precise method than the one given above for determining R_C consists in measuring the total resistance of the circuit when C is included and again when C_S has been substituted and taking the difference. The procedure followed is exactly the same as in the measurement of total resistance described in connection with the inductance coil measurements just preceding. This method is, however, considerably more laborious than the other.

INDEX